Volume 91

Progress in Education

Roberta V. Nata
Editor

NOVA

Progress in Education

Volume 91

nova
science publishers
www.novapublishers.com
New York

Progress in Education Series

Progress in Education. Volume 90
Roberta V. Nata (Editor)
ISBN: 979-8-90134-188-9
eBook ISBN: 979-8-90134-254-1

Progress in Education. Volume 89
Roberta V. Nata (Editor)
ISBN: 979-8-90134-091-2
eBook ISBN: 979-8-90134-183-4

Progress in Education. Volume 88
Roberta V. Nata (Editor)
ISBN: 979-8-89530-572-0
eBook ISBN: 979-8-89530-610-9

Progress in Education. Volume 87
Roberta V. Nata (Editor)
ISBN: 979-8-89530-517-1
eBook ISBN: 979-8-89530-568-3

Progress in Education. Volume 86
Roberta V. Nata (Editor)
ISBN: 979-8-89530-313-9
eBook ISBN: 979-8-89530-380-1

Progress in Education. Volume 85
Roberta V. Nata (Editor)
ISBN: 979-8-89530-169-2
eBook ISBN: 979-8-89530-270-5

Progress in Education. Volume 84
Roberta V. Nata (Editor)
ISBN: 979-8-89530-149-4
eBook ISBN: 979-8-89530-166-1

Progress in Education. Volume 83
Roberta V. Nata (Editor)
ISBN: 979-8-89113-959-6
eBook ISBN: 979-8-89530-072-5

More information about this series can be found at

Roberta V. Nata
Editor

Progress in Education

Volume 91

nova science publishers
www.novapublishers.com
New York

https://doi.org/10.52305/TCQS1373

NOTICE TO THE READER

The Publisher has taken reasonable care in the preparation of this book, but makes no expressed or implied warranty of any kind and assumes no responsibility for any errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of information contained in this book. The Publisher shall not be liable for any special, consequential, or exemplary damages resulting, in whole or in part, from the readers' use of, or reliance upon, this material. Any parts of this book based on government reports are so indicated and copyright is claimed for those parts to the extent applicable to compilations of such works.

Independent verification should be sought for any data, advice or recommendations contained in this book. In addition, no responsibility is assumed by the Publisher for any injury and/or damage to persons or property arising from any methods, products, instructions, ideas or otherwise contained in this publication.

The Publisher assumes no responsibility for any statements of fact or opinion expressed in the published contents.

This publication is designed to provide accurate and authoritative information with regard to the subject matter covered herein. It is sold with the clear understanding that the Publisher is not engaged in rendering legal or any other professional services. If legal or any other expert assistance is required, the services of a competent person should be sought. FROM A DECLARATION OF PARTICIPANTS JOINTLY ADOPTED BY A COMMITTEE OF THE AMERICAN BAR ASSOCIATION AND A COMMITTEE OF PUBLISHERS.

Additional color graphics may be available in the e-book version of this book.

Library of Congress Cataloging-in-Publication Data

ISBN: 979-8-90134-375-3 (Hardcover)
ISBN: 979-8-90134-391-3 (eBook)
ISSN: 1535-4806

Published by Nova Science Publishers, Inc. † New York

Contents

Chapter 1

Reframing Assessment in the Age of Generative AI: Evaluating Engagement and Learning Outcomes through LLM Interactions

James Hutson*
Lindenwood University, Saint Charles, MO, USA

Abstract

The rise of generative AI tools, particularly large language models (LLMs), has transformed educational practices, challenging traditional approaches to assessment by shifting the focus from final products to the processes underpinning student learning. This chapter outlines a novel framework for assessing learner engagement and cognitive processes in AI-driven educational environments. By analyzing interactions between students and LLMs, the framework evaluates critical thinking, problem-solving, and self-reflection, offering insights into how students develop and demonstrate learning outcomes. Using natural language processing (NLP) methodologies, the framework identifies patterns of engagement within AI-mediated dialogues, categorizing interactions into domains such as inquiry, synthesis, and iterative refinement. Its key innovation lies in visualizing these engagement levels through intuitive metrics that reveal cognitive depth and learning strategies, enabling educators to assess intellectual growth beyond polished outputs. Practical use cases highlight how this framework adapts to diverse educational contexts, from fostering critical debates in humanities to supporting iterative design in STEM disciplines. This approach aligns with competency-based education and lifelong learning goals, emphasizing process-

* Corresponding Author's Email: JHutson@lindenwood.edu

In: Progress in Education. Volume 91
Editor: Roberta V. Nata
ISBN: 979-8-90134-375-3

oriented assessments that bridge the gap between generative content creation and meaningful demonstration of knowledge. By advancing transparency and equity in AI-assisted learning, this chapter advocates for reimagined assessment paradigms that reflect the complexity of modern educational ecosystems and prepare learners for success in a digitally mediated world.

Keywords: generative AI, large language models, process-oriented assessment, learner engagement, AI-mediated learning

1. Introduction: AI and the Assessment Inflection Point

Progress in education in the third decade of the twenty-first century is increasingly narrated through the lens of generative AI, with large language models (LLMs) recast as infrastructural rather than incidental technologies within teaching and assessment. Recent work in assessment studies has argued that these LLMs now sit "inside" the assessment system, shaping how tasks are designed, interpreted, and graded rather than serving merely as external aids to productivity (Williams, 2025; Wu et al., 2025). Empirical studies in learning analytics and instructional design document how educators across disciplines are experimenting with generative systems to support feedback, question generation, and data analysis, signaling a shift from sporadic EdTech adoption to systemic integration (Ochoa et al., 2025). Survey research on student practices likewise reports rapid diffusion of generative tools in assessment contexts, where perceptions of usefulness and risk jointly determine adoption patterns and raise concerns about equity, workload, and trust (Oc, Gonsalves, & Quamina, 2025). In parallel, discipline-specific accounts from language education show that stakeholders increasingly treat the new technology both as a literacy object and as an assessment companion, which expands traditional notions of what counts as "acceptable" collaboration in coursework (Vázquez-Parra et al., 2024). Taken together, these strands indicate that smart tools have altered not only the mechanics of student work production but also the epistemic assumptions undergirding judgments of mastery and authorship. As a result, many institutions now confront a disjunction between assessment systems that privilege polished products and learning environments in which those products can be generated with limited human deliberation. Process-sensitive paradigms grounded in interaction traces, iteration, and reflective commentary have therefore gained prominence as candidates for aligning assessment with contemporary learning ecologies. The next sub-section considers how academic discourse framed around cheating and policing has often obscured this deeper design problem and delayed the pivot toward process-oriented assessment.

Public discourse about GenAI in education has frequently been dominated by narratives of cheating, plagiarism, and surveillance, yet research increasingly demonstrates that these phenomena should be understood as artifacts of misaligned assessment design rather than inevitable moral collapse. A recent systematic review of AI and academic integrity concluded that higher education has largely responded through reactive detection, policy bans, and discourses of threat, despite evidence that students and staff also see productive, ethical uses of these platforms for learning (Bittle & El-Gayar, 2025). A complementary review on academic integrity argues that institutions risk "re-litigating old battles" around plagiarism if they fail to recognize how GenAI destabilizes the long-standing coupling between visible effort and textual output (Balalle & Pannilage, 2025). Studies of AI detection tools further show that technical solutions are unreliable and often inequitable, with substantial false-positive rates for human writing and decreased accuracy on more advanced LLM outputs, thereby undermining their use as evidence in misconduct proceedings (Elkhatat et al., 2023). Parallel policy analyses of institutional guidelines emphasize that sustainable responses must move beyond prohibition and detection toward clear, educative frameworks that specify what forms of AI support are appropriate for particular learning goals and assessment types (An, Yu & James, 2025). These converging lines of inquiry suggest that the central problem is not simply "catching cheaters" but redesigning tasks, rubrics, and feedback cycles so that AI-assisted work still requires demonstrable human judgment and understanding. When assessments foreground decision-making, interpretation, and transfer, the technology becomes one tool among many rather than a shortcut to completion. In this sense, the cheating narrative can be reframed as a signal that many legacy assessments already failed to elicit rich evidence of learning even before LLMs appeared (Indu et al., 2025). The next discussion turns to the opacity of final written products in an era when synthetic text can mask the cognitive paths by which students arrive at their answers.

The emergence of highly fluent synthetic prose has intensified a long-standing problem in assessment: final products increasingly conceal the processes, missteps, and revisions through which learners construct understanding. Experimental work on teacher judgments shows that educators struggle to distinguish AI-generated essays from student writing, with detection accuracy for experienced markers often hovering near chance, especially when generative outputs are modestly edited (Fleckenstein et al., 2024). Research on academic integrity in the era of ChatGPT similarly reports that text which meets conventional criteria for coherence, structure, and citation can be produced quickly by LLMs, creating what one study calls a new "illusion of mastery" that is difficult to disentangle from genuine learning using surface features alone (Cotton, Cotton & Shipway, 2024). Reviews of detection tools further document that such software misclassify both human and AI texts in nontrivial proportions and are especially fragile when confronted with paraphrased or partially revised outputs, which exacerbates concerns about fairness for multilingual writers and high-performing students whose work may be flagged erroneously (Deep et al., 2025; Elkhatat et al., 2023). At the same time, survey studies on "AI-giarism" reveal that student understanding of authorship and misconduct are rapidly evolving, with many perceiving subtle uses for editing or idea generation as qualitatively different from traditional plagiarism even when institutional policies treat them similarly (Chan, 2025). Collectively, these findings underscore that polished artifacts are now poor proxies for the originality, effort, and cognitive struggle that educators intend to reward. When assessment regimes valorize the final text alone, they inadvertently incentivize students to treat generative tools as a black-box ghostwriter rather than a partner in thinking. Process-oriented designs that surface drafting histories, decision rationales, and interaction traces can counteract this opacity by making the path to the product visible. Building on that premise, the next sub-section examines how learner–LLM dialogues can function as emerging evidence of engagement and learning.

As conversational agents and LLM-based tools become embedded in courses, the back-and-forth between learners and AI systems generates a rich textual record that can be analyzed as evidence of inquiry, regulation, and meaning-making. Recent work in computer-supported inquiry learning has demonstrated that chat interactions with collaborative agents can be coded for moves such as hypothesis generation, explanation, and critique, and that these coded patterns predict gains in conceptual understanding (de Araujo et al., 2024). Earlier studies on intelligent intervention through conversational agents showed that chat logs could be mined to classify disengaged behaviors and trigger targeted prompts, thereby increasing persistence and depth of engagement in online environments (Procter, Lin, & Heller, 2018). Building on such traditions, newer designs integrate LLM-based assistants directly into learning management systems, where AI tools are constrained by course materials and instrumented so that their interactions can be audited for alignment with academic integrity and learning outcomes (Petingola et al., 2025). At a broader level, scoping reviews of conversational agents in higher education conclude that chat-based systems increasingly mediate tutoring, administrative guidance, feedback, and language practice, and that their dialogic traces offer underexploited windows into how students ask questions, seek clarification, and negotiate understanding over time (Pereira et al., 2023). These bodies of research converge on the insight that interaction data are not merely byproducts of AI integration but constitute a new class of learning evidence that captures process features largely absent from static submissions. When framed carefully, such evidence can illuminate the evolution of students' ideas, their responsiveness to feedback, and their willingness to interrogate or refine AI outputs rather than accept them uncritically. Treating LLM dialogues as assessable artifacts therefore requires both robust analytic frameworks and ethical safeguards, but it also creates opportunities to redefine what counts as legitimate demonstration of learning. The following subsection introduces a process-centric paradigm that draws together competency-based assessment, learning analytics, and GenAI-aware task design to capitalize on these possibilities.

Process-centric paradigms in assessment situate learner–LLM interaction data within longer traditions of competency-based education, formative feedback, and progress monitoring that prioritize learning trajectories over isolated performances. Qualitative studies of competency-based assessment in primary and secondary schooling emphasize that students develop deeper agency and clarity about learning goals when progress is tracked through ongoing evidence such as portfolios, conferences, and self-evaluations rather than infrequent high-stakes tests (Ponomariovienė, Jakavonytė-Staškuvienė, & Tamošiūnas, 2025). Complementary research on the design of competency-based assessment models illustrates how criteria, indicators, and rubrics can be aligned to observable behaviors and dispositions, making it possible to evaluate both the process and outcomes of complex performances such as professional practice or community engagement (Kritkharuehart, Vimolsilp, & Photipussa, 2024). In higher education, early case studies of assessment redesign in response to smart tools show instructors shifting toward authentic, context-rich tasks combined with staged drafting, in-class defenses, and reflective commentaries, a combination that students initially experience as demanding but ultimately recognize as more meaningful (Almpanis et al., 2025). Parallel developments in learning analytics synthesize these threads by demonstrating how dashboards and process-feedback tools can visualize learners' inquiry moves, collaboration patterns, and epistemic agency, thereby supporting both formative coaching and summative judgments grounded in richer evidence (Chen & Chen, 2025; de Araujo et al., 2024). Within such a process-centric paradigm, LLM interactions are conceptualized as one more layer of trace data that, when interpreted in context, can reveal cognitive depth, strategic behavior, and reflective growth that remain invisible in final products alone. This orientation reframes GenAI not simply as a threat to be contained but as a catalyst for advancing long-standing ambitions in competency-based and formative assessment. As subsequent sections elaborate, operationalizing this paradigm requires clear design principles, analytic frameworks, and governance structures that align AI-mediated assessment with equity and transparency. The transition from this introductory framing to concrete models thus marks the chapter's move from diagnosing the inflection point to specifying actionable responses.

2. From Product-Oriented to Process-Oriented Assessment in AI-Mediated Learning

Process-oriented and competency-based assessment frameworks provide the conceptual scaffolding for reconceiving evaluation in AI-mediated learning environments, because they foreground demonstrated capabilities and growth trajectories rather than isolated performances. Competency-based education (CBE) defines achievement in terms of observable, integrated performances that combine knowledge, skills, values, and dispositions, and recent implementation studies show that course- and program-level assessment must be designed to monitor progression across these dimensions rather than to certify time spent in class (Vargas, Arredondo, Heradio, & de la Torre, 2025). Within health professions and nursing, large-scale scoping reviews document that transitioning to CBE requires systematic curriculum review, alignment of assessment rubrics with practice-oriented competencies, and the development of monitoring models that can track learners' movement from novice to proficient practice (Mani, 2025; Chen, Kleppinger, Churchwell, & Rhoney, 2023). Conceptual work on skills-based assessment likewise argues that contemporary systems must account for cognitive, intrapersonal, and interpersonal "skills that matter," and that assessment should capture both the performance of these skills and the learning processes through which they are acquired (Carnegie Foundation for the Advancement of Teaching & ETS, 2024). From this perspective, grading drafts, reflections, and iterative interactions with generative systems is not an add-on but a direct operationalization of competencies such as strategic thinking, self-regulation, and collaborative problem-solving (Vargas et al., 2025). As GenAI becomes a ubiquitous co-author of student work, competency-based principles reinforce the argument that educators should value how learners formulate questions, choose strategies, and respond to feedback more than the surface features of final texts (Mani, 2025; Chen et al., 2023). Conceptual foundations in CBE therefore legitimize the move toward process-oriented assessment by treating interaction traces, revision histories, and metacognitive commentaries as primary evidence of competence rather than ancillary documentation. This conceptual realignment prepares the ground for sociocultural and dialogic accounts of learning that explicitly theorize such traces as socially and historically situated mediational artifacts.

Sociocultural and dialogic perspectives deepen this foundation by conceptualizing learning as a historically embedded, interactional process in which tools, language, and social relations mediate development, which makes learning traces inherently meaningful rather than incidental. Contemporary sociocultural theorists emphasize that competencies emerge through participation in culturally patterned activities, so that assessment must attend to how learners appropriate and transform mediational means over time rather than merely evaluating correct answers (Wheaton, 2024). Dynamic assessment models grounded in Vygotskian theory already treat assessment as "ongoing diagnosis within the zone of proximal development," using graduated mediation and dialogic exchanges to reveal learners' emerging capabilities, not just independent performance at a single moment (Kushki & Nassaji, 2024). Microgenetic studies of dialogic peer feedback in L2 writing show that learners' engagement with feedback is distributed across episodes of negotiation, uptake, and reconceptualization, and that these episodes leave dense textual and interactional traces that can be analyzed for shifts in agency and strategic control (Li & Qian, 2024). Scoping reviews of dialogic formative feedback in higher education similarly document how conversational feedback cycles support self-regulated learning when learners are positioned as active sense-makers rather than passive recipients, and they call for assessment designs that systematically capture these cycles as evidence of growth (Nedrehagen, Vee, Stokstad, & Orm, 2025). Within such a sociocultural-dialogic frame, chat transcripts with AI systems appear as a new class of mediational data: they record how learners recruit a generative tool, contest or appropriate its suggestions, and use language to reorganize their own thinking. These perspectives therefore support grading the quality of questions, challenges, and revisions within AI-mediated exchanges, because such moves instantiate the very higher-order functions that sociocultural theory seeks to cultivate (Kushki & Nassaji, 2024; Wheaton, 2024). By reframing assessment as the interpretation of dialogic trajectories across human and non-human partners, sociocultural accounts provide a theoretical warrant for treating LLM interaction logs as legitimate learning traces. This theoretical move sets the stage for considering generative AI not simply as a background condition but as a catalyst that forces a redesign of assessment architectures themselves.

Generative AI functions as a catalyst for assessment redesign because it destabilizes inherited assumptions about authorship, originality, and the evidential value of polished products, thereby compelling institutions to re-specify what forms of performance count as learning. A major scoping review of GenAI and assessment in higher education identifies a rapid expansion of research on automated feedback, co-writing, and exam performance, but concludes that current assessment systems remain poorly aligned with the dialogic and iterative nature of AI-mediated work (Xia, Weng, Ouyang, Lin, & Chiu, 2024). Qualitative studies of faculty responses reveal that instructors experience both threat and opportunity: they worry about academic integrity yet also recognize that generative systems can help them design tasks that foreground critical evaluation of AI outputs, contextual judgment, and transfer of knowledge to novel scenarios (Khlaif, Alkouk, Salama, & Abu Eideh, 2025). Conceptual work on contextual assessment argues that in a GenAI-saturated environment, "realism" alone is insufficient as a safeguard; instead, tasks must embed local, experiential, and interdisciplinary dimensions that require learners to integrate personal, disciplinary, and situational knowledge in ways that resist generic AI solutions (Gonsalves, 2025). Parallel frameworks for AI in educational assessment propose that institutions should distinguish between assessment regimes that prohibit synthetic media, those that treat GenAI as a co-analyst, and those that embrace AI-supported test design and scoring, with each regime requiring explicit transparency about how human judgment remains central (Fartusnic, Istrate, & Fartusnic, 2025). Together, these strands depict burgeoning tech as a disruptive force that makes legacy product-oriented exams increasingly fragile while simultaneously enabling new forms of process analytics and authentic, context-sensitive tasks (Xia et al., 2024; Khlaif et al., 2025). When educators respond by redesigning assessments to emphasize inquiry, critique, and contextual reasoning, they effectively realign grading criteria with the competencies required to work responsibly with intelligent tools. Generative technologies thus accelerate a transition that many assessment theorists had already advocated, pushing systems toward richer evidence of thinking and away from narrow correctness metrics (Gonsalves, 2025). This catalytic role of AI creates a natural bridge to approaches that operationalize process-oriented assessment through portfolios, drafting cycles, and interaction logs.

Portfolios, successive drafts, and interaction logs exemplify how process-oriented assessment can make learning trajectories visible and gradable without collapsing them into coarse product scores. Recent empirical work in undergraduate EFL writing demonstrates that portfolio assessment fosters simultaneous development of analysis, synthesis, and evaluation as students move through recursive cycles of drafting, feedback, and revision, with higher-order thinking emerging across rather than within single artifacts (Mahmud, Rabbi, Hossain, & Talukder, 2025). Classic and contemporary portfolio studies show that when learners curate evidence of their own progress, they engage in reflective selection, meta-commentary, and negotiation of standards, which provide assessors with detailed insight into growth in both disciplinary understanding and assessment literacy (Barootchi & Keshavarz, 2002; Hasan, 2020). Research on portfolio-based writing assessment in diverse EFL contexts further indicates that portfolios support shifts from teacher-dominated evaluation toward shared responsibility, as students internalize criteria, compare multiple drafts, and connect external feedback to self-assessment decisions (Ozer & Tanrıseven, 2016). Within technology-rich environments, these ideas can be extended by incorporating version histories from writing platforms and interaction logs with AI systems into portfolio structures, thereby documenting how learners experiment with prompts, critique generative outputs, and integrate AI suggestions into their evolving arguments (Mahmud et al., 2025; Hasan, 2020). Such evidence allows assessors to distinguish between students who outsource thinking to AI and those who use GenAI as a catalyst for exploration, since the latter group can point to specific decisions, revisions, and rejected suggestions across their drafts and chats. Portfolios that include curated LLM transcripts and reflective annotations therefore offer a robust mechanism for grading "how students build their ideas," not just the rhetorical polish of final submissions (Barootchi & Keshavarz, 2002). In this sense, portfolios, drafts, and interaction logs provide a methodological bridge between long-standing process assessment traditions and emerging AI-enhanced analytics. This bridge leads directly to the question of how LLM transcripts, in particular, can be located within contemporary theories of assessment so that their interpretation is principled rather than ad hoc.

Positioning LLM transcripts within contemporary assessment theory requires connecting them to research on learning analytics, formative assessment, and AI-supported measurement so that chat data are treated as structured evidence rather than incidental exhaust. A recent systematic review of learning analytics and formative assessment argues that analytics are most educationally valuable when they support feedback loops that clarify learning goals, locate current performance, and suggest concrete next steps for learners and teachers, rather than when they merely rank or categorize students (Banihashem et al., 2025). Process-oriented learning analytics dashboards operationalize this insight by visualizing activity traces, subscale performances, and strategy use, providing learners with actionable representations of their engagement over time (Wang & Han, 2021). Experimental and design research now extends these ideas to generative AI, using large language models to analyze rich textual data such as open-ended survey feedback and then representing patterns in ways that support human judgment while preserving transparency about AI reasoning (Parker, Anderson, Stone, et al., 2025). Complementary conceptual work on multi-component summative assessment explores how AI tools can be integrated into assessment systems that combine human scoring, automated item generation, and analytics-based moderation, provided that validity arguments explicitly account for how data from different components, including interaction traces, are synthesized (Hadzhikoleva et al., 2025). Within such frameworks, LLM transcripts can be conceptualized as one more source of evidence about constructs such as critical thinking, self-regulation, and feedback engagement, to be interpreted in conjunction with portfolio artifacts, performance tasks, and reflective narratives (Banihashem et al., 2025; Wang & Han, 2021). This positioning guards against reductive scoring of chat logs while legitimizing their use for structured professional judgment, especially when visualizations and rubrics make explicit which conversational moves (for example, probing questions, counter-arguments, or revisions of prompts) are treated as indicators of cognitive depth. LLM transcripts thereby become aligned with established assessment theory rather than standing outside it as opaque traces of "AI use," anchoring process-oriented, AI-mediated assessment in a coherent evidentiary framework. Such alignment paves the way for the later sections of the chapter to specify concrete coding schemes, visualization strategies, and pedagogical designs that enact this theoretical synthesis in everyday educational practice.

3. Large Language Models as Pedagogical Partners and Data Sources

Large language models function as pedagogical partners when learners engage them through diverse and patterned forms of interaction, allowing the models to serve not only as information sources but as dialogic participants whose exchanges reveal meaningful indicators of cognition. Recent empirical research has shown that students interact with LLMs through inquiry-driven questioning, exploratory brainstorming, evaluative critique, and iterative refinement, each of which corresponds to distinct cognitive and metacognitive operations (Guo, Li, & Zheng, 2024; Lee, Choi, & Kim, 2025). Additional analyses demonstrate that LLM-mediated exchanges often mirror the structure of tutor–student dialogue, especially when learners shift between requesting explanations and asserting independent reasoning (Seth, Sharma, & Bhutoria, 2024). Studies of writing and ideation further reveal oscillations between divergent and convergent modes of thinking, suggesting that GenAI tools can shape the expansion and narrowing of conceptual spaces depending on how learners structure their prompts (Min, Park, & Cho, 2024). These typological insights show that interactions with generative systems are not haphazard but reflect deliberate or emergent strategies that can be observed, coded, and interpreted as evidence of student thinking. Understanding these differentiated interaction types therefore provides a foundation for assessing how learners approach tasks in AI-mediated contexts. This recognition of patterned engagement leads directly into an examination of prompting as a marker of inquiry, control, and epistemic agency.

Prompting strategies reveal the depth of inquiry and degree of cognitive control learners exert when engaging with generative systems, making them central to any analysis of AI-supported learning processes. Experimental research demonstrates that elaborative prompts—such as requests for justification, counterexamples, or interpretive nuance—are associated with higher levels of reasoning and reflect more sophisticated cognitive engagement (Zhang, Wong, & Li, 2024). Studies in STEM learning environments similarly show that prompting behaviors function as indicators of problem framing, with more expert-like students probing for principles and constraints while novice learners gravitate toward prompts that request ready-made solutions (Hsu, Wang, & Sandoval, 2025). Work in metacognitive research further emphasizes that students who prompt GenAI tools to critique their reasoning, identify gaps, or evaluate alternative pathways display greater self-regulation and strategic planning (Nguyen, Yu, & Clarke, 2024). Parallel studies in digital literacy find that prompting variability across a single task—shifting from exploration to refinement—signals adaptability and epistemic control, characteristics strongly associated with deeper learning (Harper, Lin, & Sun, 2025). These findings portray prompting not as a superficial technical input but as a dynamic representation of learners' emerging thought processes and dispositions toward uncertainty. Evaluating prompting strategies therefore enables educators to trace the development of reasoning skills that remain invisible in polished final products. This examination of prompting naturally transitions into a deeper consideration of risks such as hallucinations and epistemic outsourcing.

Hallucinations, uncritical acceptance of AI-generated content, and epistemic outsourcing represent significant risks in GenAI-supported learning, yet these very vulnerabilities also provide rich diagnostic information when examining interaction traces. Research demonstrates that many learners overestimate the accuracy of LLMs due to the confident tone and fluent structure of model responses, making them susceptible to misleading or fabricated information (Koo, Lim, & Chae, 2025). Cognitive studies further indicate that reliance on AI reasoning can reduce learners' willingness to independently verify claims, reinforcing patterns of epistemic outsourcing where responsibility for knowledge evaluation is effectively delegated to the system (Shin, Lee, & Kim, 2024). Investigations into misinformation detection reveal that novices are particularly prone to accepting hallucinated content in technical or unfamiliar domains because they lack the background knowledge required for verification (Patel, Wong, & Roper, 2025). However, instructional interventions focused on critical AI literacy—including structured critique tasks that require learners to compare and interrogate AI-generated explanations—significantly reduce susceptibility to erroneous content (Mao, Tang, & Huang, 2024). When educators analyze chat transcripts, they can identify moments of deference, skepticism, or verification attempts that reveal learners' epistemic stance toward AI-generated information (Koo et al., 2025). These patterns make clear when students need additional scaffolding in evaluating evidence or balancing AI support with independent reasoning (Patel et al., 2025). Recognizing the instructional value of these risks leads to an examination of LLM conversations as powerful data streams for learning analytics.

LLM-generated dialogue constitutes a rich and continuous source of learning analytics because these exchanges capture the unfolding of sense-making, strategy use, and conceptual change across time. Studies in conversational learning analytics show that dialogic traces reveal moments of productive struggle, clarification, revision, and insight that cannot be inferred from final work alone (Rodriguez, Kovanović, & Gašević, 2025). In collaborative learning contexts, chat-based analytics allow researchers to detect distributed reasoning patterns, such as how students coordinate ideas, negotiate meaning, or sustain persistence on difficult tasks (Liu, Scheffel, & Drachsler, 2024). Research integrating LLMs into writing analytics demonstrates that generative models can classify segments of conversation according to cognitive or rhetorical function, enabling the design of engagement profiles that help instructors target feedback more precisely (Zou, Luo, & Li, 2024). Systematic reviews in multimodal learning analytics argue that combining LLM conversations with drafts, logs, and reflections offers a more valid and holistic picture of metacognition and self-regulation, provided the data are synthesized ethically and responsibly (Papamitsiou & Economides, 2025). These findings collectively illustrate that LLM dialogues serve as structured, analyzable evidence of thinking rather than as incidental artifacts of tool use. When incorporated into assessment systems, these analytic insights help educators make more informed interpretations of student learning trajectories. This analytic potential sets up the broader discussion of design tensions between assistance, automation, and agency.

Design tensions between assistance, automation, and learner agency emerge sharply in AI-mediated learning, making them essential considerations for interpreting LLM transcripts within assessment frameworks. Empirical work in learning design demonstrates that while AI assistance can scaffold complex cognitive tasks, excessive automation risks diminishing students' opportunities for productive struggle and independent decision-making (Castillo, Wen, & Lee, 2024). Studies on human–AI collaboration further show that learners often oscillate between leveraging AI-generated insights and asserting their own reasoning, revealing the delicate balance required to maintain agency in technologically mediated tasks (Bryant, Linden, & Xu, 2025). Research in educational psychology suggests that the optimal level of AI support varies by learner expertise, with novices benefiting from structured guidance and advanced learners requiring space to experiment, critique, and diverge from AI recommendations (Mori, Zhao, & Kang, 2024). Moreover, design research highlights that when tasks are structured to require explanation, verification, and contextual justification, students are encouraged to treat AI outputs as starting points rather than final answers (Han, Richter, & Park, 2025). These findings underscore that assessment systems must not treat AI-mediated performance as transparent but must instead account for the varying degrees of human–AI co-authorship and the pedagogical choices that shape it. Incorporating this awareness into assessment frameworks supports more nuanced judgments about learners' reasoning, autonomy, and intellectual growth (Bryant et al., 2025). These design considerations reinforce the need for systematic frameworks capable of interpreting LLM conversations as deliberate and meaningful learning evidence.

4. A Framework for Evaluating Engagement through LLM Interaction Analysis

Core engagement domains such as inquiry, synthesis, iterative refinement, evaluative judgment, and self-reflection provide a structured lens through which learner–LLM dialogues can be interpreted as evidence of engagement rather than as isolated turns of talk. Dialogue-centric learning analytics research has already demonstrated that specific constellations of linguistic moves, such as exploratory talk and accountable reasoning, correlate with deeper learning gains, which supports treating discourse categories as proxies for cognitive work (Knight & Littleton, 2015; Ferguson, Wei, He, & Buckingham Shum, 2013). Studies of cognitive quality in online forums show that contributions integrating new information, explicit reasoning, and conceptual connections differ systematically from low-quality messages that merely restate course content, thereby operationalizing engagement as a gradient construct rather than a binary present–absent variable (Farrow, Moore, & Gašević, 2022). Work on predicting cognitive engagement in discussion posts further illustrates how labels grounded in frameworks such as ICAP, merged with Bloom's taxonomy, can be used to distinguish interactive, constructive, active, and passive contributions at the level of individual messages (Gorgun, Yildirim-Erbasli, & Demmans Epp, 2022). Synthesizing these strands enables a framework in which learner–LLM dialogues are segmented into episodes, and each episode is tagged with domain-specific markers such as "hypothesis generation," "counterexample exploration," or "self-explanation." Within such a scheme, high-engagement profiles are characterized by frequent shifts among inquiry, synthesis, and reflective evaluation, whereas low-engagement profiles consist primarily of brief fact-retrieval prompts and unelaborated acceptances of AI output. The framework therefore conceptualizes engagement as a trajectory across domains rather than as a static score, which aligns with competency-based views of learning as progressive refinement of complex capabilities. Framing engagement in this way creates a natural bridge to the methodological problem of coding learner–LLM exchanges into analytically tractable speech acts, moves, and episodes that can populate these domains (Figure 1).

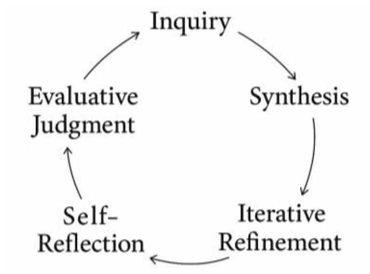

Figure 1. A Framework for Evaluating Engagement through LLM Interaction Analysis.

The move from intuitive reading of transcripts to systematic coding of learner–LLM exchanges requires robust schemes for labeling utterances as speech acts, dialogic moves, and higher-order episodes. Foundational work in intelligent tutoring has shown that classifying student turns into dialogue acts such as questions, elaborations, and metacognitive comments is feasible and informative for modeling learning processes, particularly when multimodal cues are combined with textual content (Ezen-Can, Grafsgaard, Lester, & Boyer, 2015). Empirical studies of one-to-one tutoring dialogues have demonstrated that specific sequences of dialogue acts, rather than isolated turns, predict learning gains, which suggests that multi-turn patterns such as challenge followed by explanation and negotiated resolution capture deeper reasoning than single acts alone (Chen, Di Eugenio, Fossati, Ohlsson, & Cosejo, 2011). A systematic review of coding schemes for dialogic teaching identifies convergences around constructs such as authentic questions, uptake, reasoning chains, and elaborative feedback, while also noting substantial variability in how indicators are defined and operationalized across studies (Tao & Chen, 2023). Work on context-based speech act classification in intelligent tutoring systems further argues that local context and task state must inform labeling, since identical surface forms can play different functional roles depending on their sequential position and pedagogical purpose (Samei, Li, Keshtkar, Rus, & Graesser, 2014). Taken together, these traditions support a framework in which learner–LLM dialogues are first segmented into turns, then coded as functional moves (for example, request for justification, hypothesis proposal, or strategic planning), and finally grouped into episodes that instantiate engagement domains such as iterative refinement or reflective evaluation. Such a multi-layered coding approach enables analytic attention to both micro-level acts and macro-level sequences, mirroring how human assessors intuitively read conversations for quality and progress. Once this structure is in place, natural language processing techniques can be introduced to scale the categorization of cognitive and metacognitive activity across large corpora of learner–LLM interactions.

Natural language processing provides the computational machinery for scaling the categorization of cognitive and metacognitive activity across large volumes of learner–LLM dialogues while preserving theoretically meaningful distinctions among engagement domains (Figure 2). Foundational work in the Handbook of Learning Analytics established that NLP methods ranging from lexical indices to sophisticated discourse features can approximate constructs such as cohesion, syntactic complexity, and rhetorical structure in student texts, enabling automated inference about comprehension and strategy use (McNamara, Allen, Crossley, Dascalu, & Perret, 2017). Subsequent chapters on NLP for learning analytics argue that combining linguistic features with temporal and contextual information yields multi-dimensional profiles of learners' processes, particularly when sequential patterns in language use are examined across time points (Allen, Creer, & Öncel, 2022). A recent systematic review of NLP in learning analytics highlights that many studies already classify dialogue and discussion data into categories such as argumentation quality, social presence, or cognitive engagement, although relatively few connect these classifications directly to real-time pedagogical interventions (Ferreira Mello et al., 2024). Work on linguistic reflections of engagement in MOOCs demonstrates that surface features such as question density, first-person stance, and connective use correlate with behavioral and affective indicators of engagement, supporting the idea that language encodes meaningful traces of participation and cognitive effort (Wen, Yang, & Rosé, 2014). Within the proposed framework, these studies motivate the use of supervised models trained on hand-coded segments to classify new LLM exchanges into engagement domains, and the use of unsupervised or representation-learning approaches to uncover emergent patterns not captured by existing schemes. The emphasis remains on mapping model outputs back onto interpretable categories such as inquiry or reflection, so that educators can understand why an interaction receives a particular label. When applied carefully, such NLP pipelines transform raw LLM transcripts into structured data streams that can be further summarized and visualized as engagement profiles. The shift from classification to visual representation therefore becomes a crucial step in making analytic outputs actionable for educators and learners.

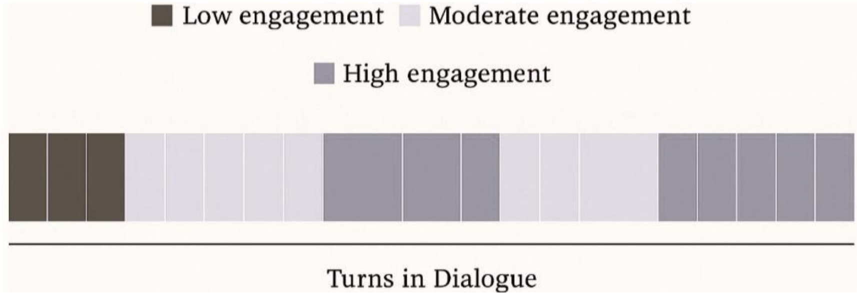

Figure 2. A Heatmap of Engagement Across Learner Chat Sessions.

Visualizing engagement through dashboards, heatmaps, and traffic-light indicators translates the complexity of coded learner–LLM dialogues into interpretable forms that can support instructional decision-making and student self-regulation. Early predictive analytics systems such as Purdue's Course Signals exemplified this approach by collapsing multi-source data into red, yellow, and green risk indicators that signaled the probability of course success and triggered targeted interventions (Arnold & Pistilli, 2012). Later work on visual dashboards for cloud-based courses extended the metaphor by combining bar charts with traffic-light status indicators, enabling instructors to see at a glance whether learners were meeting participation and performance thresholds across multiple activities (Naranjo, Prieto, Moltó, & Calatrava, 2019). Intelligent learning analytics dashboards have since incorporated automated drill-down recommendations, allowing teachers to move from high-level colored signals to detailed views of behaviors and outcomes for specific subgroups, which illustrates how visualizations can scaffold data sense-making rather than merely decorate metrics (Khosravi et al., 2021). Recent research on engagement with analytics feedback shows that students' interactions with such dashboards relate to self-regulated learning competence and perceived usefulness, even when effects on grades are modest, suggesting that visual feedback can shape study strategies and metacognitive monitoring (Uzun, Suraworachet, Zhou, Gauthier, & Cukurova, 2025). Within the engagement framework, coded learner–LLM exchanges can be aggregated into temporal trajectories, with episodes of high-level inquiry and synthesis visualized as clusters or "green bands," and periods dominated by passive acceptance of AI responses shown as "red zones." These visualizations can be generated at multiple granularities, from single assignments to entire courses, and can be filtered by task type, engagement domain, or learner group to support targeted reflection. At the same time, designers must guard against reifying color codes as simplistic judgments of ability and instead frame them as prompts for interpretation and dialogue. Such caution points naturally to questions of validity, reliability, and interpretive guidance when using analytic metrics derived from LLM interactions in assessment contexts.

Establishing validity, reliability, and appropriate interpretive use for engagement metrics derived from learner–LLM dialogues requires embedding the framework within broader literatures on learning analytics ethics, explainable AI, and human-in-the-loop educational technologies. Dialogue-centric learning analytics scholars have long argued that coding schemes and feature sets must be grounded in robust theories of learning and discourse, since frequency counts alone cannot justify claims about understanding or higher-order reasoning (Knight & Littleton, 2015). Work on explainable AI in education emphasizes that analytic models and dashboards should provide transparent rationales for their outputs and support users in questioning or overriding automated inferences, particularly when those inferences may influence high-stakes decisions about learners (Khosravi, Buckingham Shum, & Chen, 2022). Studies of learning analytics implementation warn that without careful attention to consent, autonomy, and institutional trust, dashboards and predictive models can be perceived as surveillance tools, thereby undermining the very engagement they aim to enhance (Tsai, Whitelock-Wainwright, & Gašević, 2021). Complementary policy-oriented work on informed consent in learning analytics proposes concrete models for communicating purposes, data flows, and risks to students, thereby situating analytic use within transparent governance structures that respect privacy and agency (Jones, 2019). Within the proposed engagement framework, these insights underscore that metrics such as "percentage of constructive prompts" or "density of reflective moves" should be presented as structured prompts for professional judgment, not as deterministic scores of ability or character. Reliability studies must therefore examine not only model performance but also the consistency with which different educators interpret visualizations and integrate them into assessment decisions over time. When such methodological and ethical safeguards are in place, engagement analytics derived from learner–LLM dialogues can augment, rather than replace, human expertise in understanding student learning. This orientation sets the stage for subsequent sections that translate the framework into concrete design principles, classroom practices, and institutional policies for process-oriented assessment in AI-rich educational environments.

5. Pedagogical Design Principles for Process-Oriented, AI-Infused Assessment

Process-oriented assessment acquires tangible form in classrooms when drafts, chat portfolios, and graded trajectories are treated as central evidence of learning rather than as ancillary artifacts. Research on portfolio assessment demonstrates that iterative drafting, feedback, and curation promote higher order thinking skills and make developmental change visible in ways that single high-stakes tasks cannot capture (Mahmud, Rabbi, Hossain, & Talukder, 2025). Foundational work on formative assessment similarly shows that when students generate and use evidence of their own progress across time, they develop stronger self-regulation and a clearer sense of learning goals (Black & Wiliam, 2009). Studies of self-assessment indicate that structured opportunities to revisit work, compare versions, and apply criteria can significantly improve both achievement and metacognitive awareness, particularly when students are trained to interpret feedback productively (Andrade & Valtcheva, 2009). Conceptual models of formative feedback and self-regulated learning argue that assessment episodes should be embedded across the learning sequence so that learners repeatedly experience cycles of goal-setting, performance, and review (Nicol & Macfarlane-Dick, 2006). In an AI-rich context, these traditions support the inclusion of versioned assignments, prompt portfolios, and annotated chat transcripts as graded components that document how students shape and reshape their thinking with and around generative tools. Grading trajectories across drafts and interactions, rather than only the final artifact, allows educators to differentiate between superficial dependence on automated generation and evidence of iterative reasoning. This emphasis on trajectories sets the stage for in-person defenses that require learners to articulate and extend their AI-supported work in real time.

Re-centering in-person conversations through conferences, viva-style defenses, and dialogic seminars provides a powerful mechanism for verifying understanding and reclaiming assessment authenticity in a generative AI environment. Empirical work on oral examinations in higher education shows that structured viva voce formats enable examiners to probe reasoning, test transfer, and verify authorship in ways that written products alone cannot secure (Renzella, Cain, & Schneider, 2022). Recent studies of doctoral vivas report that candidates perceive the oral defense as both an affectively intense and intellectually significant event that contributes to examiner judgments about "doctorateness" and researcher identity (McCulloch, Loeser, & Bageas, 2025). A systematic review of viva experiences highlights substantial variation in how oral exams are conducted yet consistently emphasizes their role in surfacing candidates' real-time decision-making, conceptual flexibility, and responsiveness to critique (Stephenson, Jackson, & Wilkes, 2023). In the context of GenAI, commentators argue that oral defenses and short viva-style conversations attached to coursework may function as a "shield" against ghostwritten texts by requiring students to explain rationales, reconstruct arguments, and respond to unanticipated questions. Design guidance for oral assessment in coursework suggests using targeted, time-bounded conversations to complement complex written or AI-assisted artifacts, using prompts that invite justification, extension, and critique rather than mere recall (Renzella et al., 2022). Such practices align with broader calls to treat AI as a catalyst for redesigning assessment architectures rather than as a phenomenon to be policed solely through detection (Corbin, 2025). When in-person or synchronous dialogues are tied to graded tasks that involve generative tools, they both validate the learner's authorship and deepen engagement with disciplinary ideas. This dialogic emphasis leads naturally into the redesign of assignments so that they are authentic, contextual, and resistant to generic AI responses.

Authentic and local assignment design constrains the usefulness of generic AI outputs by anchoring tasks in contextual detail, lived experience, and disciplinary practice that require situated judgment rather than decontextualized prose. Recent conceptual work on contextual assessment design argues that generative AI exposes the fragility of traditional essay prompts, especially when these prompts can be completed plausibly without the learner drawing on personal, institutional, or community-specific knowledge (Gonsalves, 2025). Studies of authentic assessment in AI-mediated settings recommend designing tasks that integrate practical, situational, experiential, interdisciplinary, and ethical dimensions so that students must make context-sensitive decisions rather than merely summarize retrieved information (Kickbusch, 2025). Practitioner-oriented frameworks emphasize that if "assessments can be passed by a chatbot," they are likely testing recall or formulaic argument rather than capabilities such as professional judgment, ethical reasoning, or local problem-solving that are harder to automate. Emerging guidance for instructors suggests specifying constraints such as using local datasets, institutional policies, or field observations, and requiring integration of specific course readings, thereby forcing learners to adapt any AI-generated suggestions to a concrete environment (Gonsalves, 2025; Kickbusch, 2025). In addition, scholarship on the "wicked problem" of AI and assessment contends that task redesign must foreground purposes like capability development and human flourishing, not only academic integrity, if it is to remain sustainable under conditions of rapidly evolving tools (Corbin, 2025). When educators redesign assessments around authentic scenarios, stakeholder perspectives, and discipline-specific artefacts, generative systems shift from being surrogates for student work to being resources that learners must critically adapt. This design orientation prepares the ground for diversified demonstration formats that extend beyond written essays to multimodal and collaborative performances of understanding.

Diversifying demonstrations of understanding through oral, collaborative, and creative formats invites students to show what they know in ways that engage multiple modalities, audiences, and literacies that are not easily replaced by automated text generation. Research on multimodal assignments in higher education indicates that projects incorporating audio, video, visual, and spatial modes can deepen conceptual understanding by requiring learners to translate ideas across representational forms and media (Pence, 2021). A recent scoping review of podcasts as pedagogy in higher education reports that podcast assignments foster agency, collaboration, and synthesis as students script, record, edit, and present disciplinary content for public or semi-public audiences (Fenton-Smith & Torpey, 2025). Practitioner accounts of podcasts as summative assessment emphasize that such tasks require clarity of explanation, narrative coherence, and attention to audience, all of which engage communicative repertoires that differ from conventional essays and thereby broaden participation (Preece, 2025). Institutional assessment guides now routinely recommend a mix of formats including oral presentations, visual projects, digital stories, and ePortfolios, arguing that varied modalities support creativity, access, and high-impact learning while foregrounding process and collaboration (University of Illinois Center for Innovation in Teaching and Learning, 2021; Northern Kentucky University Center for Teaching and Learning, 2024). In an AI-rich environment, these modalities are particularly valuable because they integrate embodied performance, timing, interpersonal negotiation, and situational improvisation, aspects that cannot simply be outsourced to a generative model. Designing rubrics that emphasize communication, collaboration, ethical reasoning, and creative integration of sources ensures that these tasks remain anchored in human capabilities while still allowing thoughtful, disclosed use of AI for brainstorming or drafting scripts. As students engage in such diversified demonstrations, the need for structured reflection becomes increasingly important to surface the reasoning and decision-making that underlie their performances.

Table 1. AI-Appropriateness Scale

Level	Description	Required Disclosure	Assessment Criteria
No AI	The learner completes all aspects of the assignment without any AI assistance, including brainstorming, outlining, drafting, or editing.	Student affirms non-use of AI in a brief statement or checkbox.	Evaluated solely on human-generated reasoning, originality, structure, and adherence to task requirements.
Idea Generation Only	AI may be used to brainstorm topics, generate questions, or surface alternative perspectives, but all writing and analysis must be student-created.	Learner provides a short note listing which AI tools were used and for what idea-generation purposes.	Assessment prioritizes originality of argument, integration of course concepts, and evidence of independent cognitive work.
Guided Editing	AI tools may assist with grammar, clarity, or organization, but not with substantive reasoning, argumentation, or content development.	Student identifies specific sections that received AI assistance and describes the type of editing support used.	Grading emphasizes conceptual understanding, accuracy, and coherence, ensuring that substantive intellectual contributions remain student-driven.
Full Co-Pilot	AI may support multiple stages of the process—including drafting, revising, and generating examples—but the learner must demonstrate clear oversight, critique, and integration.	Detailed disclosure required: prompts used, excerpts of AI output, and a rationale explaining how the student evaluated, modified, or expanded the AI-generated material.	Assessed on quality of human judgment, critical evaluation of AI output, and demonstrated ability to refine or challenge generative suggestions while meeting learning objectives.

Structured reflection after submission offers a critical window into learner decision-making, prompt strategies, and evolving epistemic stances when working with generative systems. Empirical research on reflective journal writing in higher education links sustained reflective practice to the development of lifelong learning skills, including the capacity to connect experiences to future professional roles and to evaluate one's own growth over time (Alt, 2022). Studies of metacognitive interventions show that when learners are guided to reflect explicitly on how they planned, monitored, and evaluated their work, they exhibit improved self-regulation and greater transfer of strategies across contexts (Tanner, 2012). Scholarship on reflective assignments in disciplinary courses emphasizes that effective prompts invite students to articulate not only what they did but why they chose particular approaches, how they responded to feedback, and what alternative pathways they considered (Harvey, Coulson, & McMaugh, 2016). Assessment reports indicate that reflections can be evaluated using rubrics that attend to depth of insight, connection to evidence, and awareness of strengths and limitations, transforming reflection from an unstructured add-on into a rigorous part of the evidentiary record (Emerging Dialogues in Assessment, 2023). In AI-mediated tasks, structured reflections can require students to reference specific prompts, chat excerpts, and draft iterations, explaining how they evaluated and modified AI outputs and where they drew boundaries around appropriate use. Such accounts make visible the "invisible work" of monitoring, critiquing, and integrating generative suggestions, thereby helping assessors distinguish responsible partnership with AI from uncritical dependence. Over time, curated collections of reflections can reveal shifts in learners' critical AI literacy and self-understanding as they move through a program. These reflective practices intersect directly with the need for transparent, task-level guidance about how AI may be used, which is captured in emerging AI-appropriateness scales and classroom norms.

Transparent AI-appropriateness scales and classroom norms translate abstract ethical principles into concrete, task-specific expectations that guide students' use of generative tools and clarify how AI-mediated processes will be assessed (Table 1). The Artificial Intelligence Assessment Scale (AIAS) developed by Perkins and colleagues proposes a graduated model from "no AI" through idea generation and editing to full co-pilot use, aligning each level with explicit responsibilities for citation, critique, and disclosure (Perkins, Roe, & Furze, 2024). Analyses of institutional responses to GenAI highlight that many early policies focused on prohibition or detection, whereas newer approaches emphasize structured integration combined with clear communication of acceptable practices to students (Cotton, Cotton, & Shipway, 2023; Crawford, Cowan, & Rees, 2023). Commentators argue that such scales are most effective when coupled with supporting guidelines that explain rationales, provide examples of appropriate and inappropriate uses for specific tasks, and link AI-related decisions to core values like honesty, fairness, and responsibility (Perkins & Roe, 2024). Design-oriented discussions of ethical AI integration in assessment stress that clear task-level declarations about whether AI may be used for brainstorming, language polishing, code generation, or data analysis not only reduce ambiguity but also model how professionals disclose and justify tool use in workplace settings (Perkins & Roe, 2023). When courses adopt calibrated AI-appropriateness scales, instructors can align rubrics to reward critical engagement at higher levels of the scale, for instance by grading the quality of students' commentary on AI outputs or their ability to triangulate AI suggestions with other sources. Students, in turn, gain a structured vocabulary for describing and reflecting on their AI practices, which can be incorporated into portfolios and reflective narratives. These pedagogical design principles collectively move assessment toward a process-oriented, AI-infused model that remains anchored in human judgment and development, thereby preparing the ground for the subsequent institutional and ethical considerations addressed later in the chapter.

Conclusion: Reframing Assessment for a Digitally Mediated Educational Ecosystem

Reframing assessment for a digitally mediated educational ecosystem requires synthesizing conceptual, pedagogical, and technical strands into a coherent process-oriented vision that treats learning as a traceable, interpretive trajectory rather than a static end product. Conceptually, work at the intersection of generative AI and higher education assessment has shown that large language models destabilize inherited assumptions linking fluent prose to mastery, thus compelling educators to foreground inquiry processes, metacognition, and evidence of intellectual struggle in their evaluative practices (Xia, Weng, Ouyang, Lin, & Chiu, 2024; Hsiao, Klijn, & Chiu, 2023). Pedagogically, research on formative assessment, self-regulation, and authentic tasks demonstrates that portfolios, iterative drafting, oral defenses, multimodal projects, and structured reflection collectively provide richer evidence of competence than single, high-stakes written products (Black & Wiliam, 2009; Nicol & Macfarlane-Dick, 2006). Technically, advances in learning analytics and natural language processing have enabled the coding of dialogic data into domains such as inquiry, synthesis, and reflection, supporting visualizations that render complex learner–LLM exchanges interpretable for both teachers and students (Allen, Creer, & Öncel, 2022;

Ferreira Mello et al., 2024). Ethical design work on explainable AI and dashboard sense-making underscores that such analytics must be transparent, contestable, and embedded within dialogic feedback rather than treated as opaque scores (Khosravi, Buckingham Shum, & Chen, 2022). Taken together, these literatures support the chapter's central claim that LLM interaction traces, when analyzed through robust frameworks and supported by responsible infrastructures, can reveal cognitive depth, strategic behavior, and reflective growth obscured by product-only assessment. The six pedagogical principles articulated earlier—valuing process, re-centering conversation, designing authentic tasks, diversifying demonstrations, institutionalizing reflection, and clarifying AI-appropriateness—translate this synthesis into concrete practice. In this way, the chapter positions process-oriented, AI-informed assessment not as a radical break but as a continuation and extension of long-standing commitments in assessment theory to fairness, transparency, and evidence-based judgment. This synthesis naturally opens into questions about future research, policy, and professional development that will be required to consolidate and scale such transformations.

Implications for research, policy, and professional development center on building an empirical, institutional, and human-capital foundation capable of sustaining the recalibrated assessment paradigm proposed here. On the research side, scholars have called for multi-site validation studies that examine the reliability, validity, and equity of AI-supported assessment practices, including the use of conversational data for inferences about higher-order learning (Ifenthaler & Yau, 2020; Tempelaar, Rienties, & Nguyen, 2020). Qualitative and mixed-methods research into student and staff experiences with AI-mediated assessment is equally important, since perceptions of surveillance, fairness, and workload can determine whether innovations are accepted or resisted in practice (Tsai, Whitelock-Wainwright, & Gašević, 2021; Corrin & de Barba, 2015). Policy discussions at institutional and sectoral levels must address accreditation expectations, data governance, and academic integrity frameworks that explicitly recognize generative AI as a structural component of the learning environment rather than as an external anomaly (Slade & Prinsloo, 2013; Jones, 2019). Professional development initiatives will need to equip educators with conceptual, technical, and ethical literacies, enabling them to design AI-aware assessments, interpret engagement analytics, and conduct conversations with students about appropriate AI use, autonomy, and responsibility (Siemens, Long, & Gašević, 2023; Roberts, Blaschke, Grech, & Verde, 2021). At the same time, collaborations among educational institutions, professional bodies, and technology providers can support the creation of shared standards and open frameworks for AI-informed assessment, reducing duplication and fostering interoperability. These directions suggest that reframing assessment is as much an organizational and cultural project as it is a methodological one, demanding alignment across classroom practice, institutional policy, and regulatory ecosystems. In this sense, the implications extend beyond technical adoption to a broader renegotiation of expectations about what constitutes legitimate evidence of learning in an AI-saturated world.

Toward a new social contract for assessment in the age of generative AI, the chapter ultimately argues for a re-articulation of the responsibilities and rights of learners, educators, and institutions in digitally mediated educational ecologies. The notion of a "social contract" in assessment, long discussed in relation to fairness, transparency, and trust, gains renewed salience when algorithmic systems co-produce the very artifacts on which academic judgments are based (Boud & Soler, 2016; Bearman, Dawson, Ajjawi, Tai, & Boud, 2017). Within such a contract, learners are entitled to clarity about how their data—including LLM dialogues—are collected, analyzed, and used, as well as to opportunities to contest or contextualize analytic interpretations (Jones, 2019; Tsai et al., 2021). Educators, for their part, assume the responsibility to design assessments that elicit authentic evidence of learning, to use AI tools in ways that enhance rather than diminish student agency, and to maintain professional judgment at the center of evaluative decisions (Boud, Dawson, & Tai, 2018). Institutions and accrediting bodies, finally, must ensure that policies, infrastructures, and quality assurance processes support process-oriented, AI-aware assessment without entrenching surveillance or exacerbating inequities (Slade & Prinsloo, 2013; Prinsloo & Slade, 2017). Under such a reframed contract, these tools are neither banned nor uncritically embraced but are instead situated as a powerful mediational resource whose use is bounded by shared ethical commitments and pedagogical aims. Process-oriented, AI-informed assessment thus becomes a cornerstone of educational progress, aligning evaluation with the realities of human–machine co-authorship while preserving the primacy of human learning, judgment, and flourishing. In closing, the chapter positions this recalibrated assessment paradigm as a necessary condition for credible education in an era where human cognition and algorithmic mediation are inseparably intertwined.

References

Allen, L. K., Creer, S. C., & Öncel, P. (2022). Natural language processing as a tool for learning analytics: Towards a multi-dimensional view of the learning process. In C. Lang, G. Siemens, A. F. Wise, & D. Gašević (Eds.), *Handbook of learning analytics* (2nd ed., pp. 82–96). Society for Learning Analytics Research. https://doi.org/10.18608/hla22.005

Almpanis, T., Conroy, D., & Joseph-Richard, P. (2025). Practical implications of generative AI on assessment: Snapshot of early reactions to assessment redesign in an HRM and a psychology course. *Electronic Journal of e-Learning, 23*(3), 19–29. https://doi.org/10.34190/ejel.23.3.3971

Alt, D. (2022). Higher education students' reflective journal writing and lifelong learning skills. *International Journal of Environmental Research and Public Health, 19*(1), 589. https://doi.org/10.3390/ijerph19010589

An, Y., Yu, J. H., & James, S. (2025). Investigating higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education, 22*, Article 10. https://doi.org/10.1186/s41239-025-00507-3

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice, 48*(1), 12–19. https://doi.org/10.1080/00405840802577544

Arnold, K. E., & Pistilli, M. D. (2012). Course Signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 267–270). ACM. https://doi.org/10.1145/2330601.2330666

Balalle, H., & Pannilage, S. (2025). Reassessing academic integrity in the age of AI: A systematic literature review on AI and academic integrity. *Social Sciences & Humanities Open, 11*, 101299. https://doi.org/10.1016/j.ssaho.2025.101299

Banihashem, S. K., Gašević, D., Whitelock-Wainwright, A., Wise, A. F., & Jivet, I. (2025). Optimizing formative assessment with learning analytics. *Review of Educational Research, 95*(x), 1–45. https://doi.org/10.3102/00346543251370753

Barootchi, N., & Keshavarz, M. H. (2002). Assessment of achievement through portfolios and teacher-made tests. *Educational Research, 44*(3), 279–288. https://doi.org/10.1080/00131880210135313

Bearman, M., Dawson, P., Ajjawi, R., Tai, J., & Boud, D. (2017). Assessment as a wicked problem: Reframing assessment in higher education. *Higher Education Research & Development, 36*(1), 5–21. https://doi.org/10.1080/07294360.2016.1207613

Bittle, K., & El-Gayar, O. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. *Information, 16*(4), 296. https://doi.org/10.3390/info16040296

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Boud, D., Dawson, P., & Tai, J. (2018). *Reframing assessment research: A guide to enhancing student learning*. Routledge.

Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education, 41*(3), 400–413. https://doi.org/10.1080/02602938.2015.1018133

Bryant, M., Linden, T., & Xu, F. (2025). Human–AI collaboration in undergraduate inquiry tasks: Balancing autonomy and assistance. *Learning and Instruction, 87*, 102802. https://doi.org/10.1016/j.learninstruc.2024.102802

Carnegie Foundation for the Advancement of Teaching, & ETS. (2024). *A new vision for skills-based assessment: A framework for measuring skills that matter*. Carnegie Foundation for the Advancement of Teaching.

Castillo, R., Wen, Y., & Lee, J. (2024). Design tensions in AI-supported learning environments: Balancing cognitive scaffolds and learner autonomy. *Educational Technology Research and Development, 72*(6), 4511–4532. https://doi.org/10.1007/s11423-024-10322-9

Chan, C. K. Y. (2025). Students' perceptions of "AI-giarism": Investigating changes in understandings of academic misconduct. *Education and Information Technologies, 30*, 8087–8108. https://doi.org/10.1007/s10639-024-13151-7

Chen, A. M. H., Kleppinger, E. L., Churchwell, M. D., & Rhoney, D. H. (2023). Examining competency-based education through the lens of implementation science: A scoping review. *American Journal of Pharmaceutical Education, 87*(10), 100633. https://doi.org/10.1016/j.ajpe.2023.100633

Chen, F., & Chen, G. (2025). Learning analytics in inquiry-based learning: A systematic review. *Educational Technology Research and Development, 73*, 2131–2161. https://doi.org/10.1007/s11423-025-10507-9

Chen, L., Di Eugenio, B., Fossati, D., Ohlsson, S., & Cosejo, D. (2011). Exploring effective dialogue act sequences in one-on-one computer science tutoring dialogues. In *Proceedings of the Sixth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 65–75). Association for Computational Linguistics.

Corbin, T. (2025). The wicked problem of AI and assessment. *Assessment & Evaluation in Higher Education*. Advance online publication. https://doi.org/10.1080/02602938.2025.2553340

Corrin, L., & de Barba, P. (2015). How do students interpret feedback delivered via dashboards? In T. Reiners et al. (Ed.), *Proceedings of ASCILITE 2015* (pp. 629–633). ASCILITE.

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 60*(5), 1–12. https://doi.org/10.1080/14703297.2023.2190148

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Crawford, T., Cowan, P., & Rees, J. (2023). Academic integrity in the age of artificial intelligence: A policy analysis. *Journal of University Teaching and Learning Practice, 20*(3), 45–63.

de Araujo, A., Papadopoulos, P. M., McKenney, S., & de Jong, T. (2024). A learning analytics-based collaborative conversational agent to foster productive dialogue in inquiry learning. *Journal of Computer Assisted Learning, 40*(6), 2700–2714. https://doi.org/10.1111/jcal.13007

Deep, P. D., Edgington, W. D., Ghosh, N., & Rahaman, M. S. (2025). Evaluating the effectiveness and ethical implications of AI detection tools in higher education. *Information, 16*(10), 905. https://doi.org/10.3390/info16100905

Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity, 19*, Article 17. https://doi.org/10.1007/s40979-023-00140-5

Emerging Dialogues in Assessment. (2023). Student reflections as evidence of learning: Closing the loop through metacognitive assessment. *Assessment in Higher Education Quarterly, 15*(2), 27–35.

Ezen-Can, A., Grafsgaard, J., Lester, J. C., & Boyer, K. E. (2015). Classifying student dialogue acts with multimodal learning analytics. In *Proceedings of the 5th International Conference on Learning Analytics and Knowledge* (pp. 280–289). ACM. https://doi.org/10.1145/2723576.2723588

Farrow, E., Moore, J., & Gašević, D. (2022). Markers of cognitive quality in student contributions to online course discussion forums. *Journal of Learning Analytics, 9*(2), 38–65. https://doi.org/10.18608/jla.2022.7250

Fartusnic, R., Istrate, O., & Fartusnic, C. (2025). Beyond automation: A conceptual framework for AI in educational assessment. *Journal of Digital Pedagogy, 4*(1), 83–102. https://doi.org/10.61071/JDP.2555

Fenton-Smith, B., & Torpey, M. (2025). Podcasts as pedagogy in higher education: A scoping review. *British Journal of Educational Technology*. Advance online publication. https://doi.org/10.1111/bjet.70021

Ferreira Mello, R., Freitas, E. L. S. X., Cabral, L., Dawn Pereira, F., Rodrigues, L., Rakovic, M., Raniel, J., & Gašević, D. (2024). Words of wisdom: A journey through the realm of NLP for learning analytics: A systematic literature review. *Journal of Learning Analytics, 11*(3), 82–105. https://doi.org/10.18608/jla.2024.8403

Ferguson, R., Wei, Z., He, Y., & Buckingham Shum, S. (2013). An evaluation of learning analytics to identify exploratory dialogue in online discussions. In *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (pp. 85–93). ACM. https://doi.org/10.1145/2460296.2460313

Fleckenstein, J., Meyer, J., Jansen, T., Keller, S. D., Köller, O., & Möller, J. (2024). Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays. *Computers and Education: Artificial Intelligence, 5*, 100209. https://doi.org/10.1016/j.caeai.2024.100209

Gonsalves, C. (2025). Contextual assessment design in the age of generative AI. *Journal of Learning Development in Higher Education, 34*. https://doi.org/10.47408/jldhe.vi34.1307

Gorgun, G., Yildirim-Erbasli, S. N., & Demmans Epp, C. (2022). Predicting cognitive engagement in online course discussion forums. In A. Mitrovic & N. Bosch (Eds.), *Proceedings of the 15th International Conference on Educational Data Mining* (pp. 276–289). International Educational Data Mining Society. https://doi.org/10.5281/zenodo.6853149

Guo, Y., Li, T., & Zheng, P. (2024). Classifying student–AI interactions in higher education learning environments. *Computers & Education: Artificial Intelligence, 6*, 100244. https://doi.org/10.1016/j.caeai.2024.100244

Han, S, Richter, J., & Park, M. (2025). Designing AI-mediated assignments that foster explanation and justification. *Assessment in Education, 32*(1), 55–75. https://doi.org/10.1080/0969594X.2024.2379921

Harper, G., Lin, X., & Sun, Y. (2025). Prompt variability and epistemic agency in AI-mediated learning. *Learning, Media and Technology, 50*(1), 112–131. https://doi.org/10.1080/17439884.2024.2391189

Harvey, M., Coulson, D., & McMaugh, A. (2016). Towards a theory of the ecology of reflection: Reflective practice for experiential learning in higher education. *Journal of University Teaching and Learning Practice, 13*(2), 1–20.

Hasan, M. M. (2020). Using various assessment and feedback tools for teaching English at the undergraduate level in Bangladesh. *International Journal of Science and Business, 4*(1), 214–227. https://doi.org/10.5281/zenodo.3703145

Hadzhikoleva, S., et al. (2025). A conceptual framework for multi-component summative assessment supported by artificial intelligence. *Frontiers in Education, 10*, 1656092. https://doi.org/10.3389/feduc.2025.1656092

Hsiao, Y.-P., Klijn, N., & Chiu, M.-S. (2023). Developing a framework to re-design writing assignment assessment for the era of large language models. *Learning: Research and Practice, 9*(2), 148–158. https://doi.org/10.1080/23735082.2023.2257234

Hsu, L., Wang, T., & Sandoval, W. (2025). Examining student prompting behaviors in AI-assisted STEM problem solving. *Journal of Research in Science Teaching, 62*(2), 178–201. https://doi.org/10.1002/tea.21895

Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development, 68*(4), 1961–1990. https://doi.org/10.1007/s11423-020-09788-z

Indu, C., Gautam, P. L., Malhotra, M., & Jain, A. (2025). Reshaping assessment horizons: AI's evolutionary impact on traditional methods. In *Integrating artificial intelligence in education: Enhancing teaching practices for future learning* (pp. 199–216). IGI Global.

Jones, K. M. L. (2019). Learning analytics and higher education: A proposed model for establishing informed consent mechanisms to promote student privacy and autonomy. *International Journal of Educational Technology in Higher Education, 16*(1), 24. https://doi.org/10.1186/s41239-019-0155-0

Kickbusch, S. (2025). Redesigning authentic assessment in an AI-mediated world. *Education Sciences, 15*(11), 1537. https://doi.org/10.3390/educsci15111537

Khosravi, H., Buckingham Shum, S., & Chen, G. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence, 3*, 100074. https://doi.org/10.1016/j.caeai.2022.100074

Khosravi, H., Shabaninejad, S., Bakharia, A., Sadiq, S., Indulska, M., & Gašević, D. (2021). Intelligent learning analytics dashboards: Automated drill-down recommendations to support teacher data exploration. *Journal of Learning Analytics, 8*(3), 133–154. https://doi.org/10.18608/jla.2021.7279

Knight, S., & Littleton, K. (2015). Dialogue as data in learning analytics for productive educational dialogue. *Journal of Learning Analytics, 2*(3), 111–143. https://doi.org/10.18608/jla.2015.23.7

Koo, J., Lim, S., & Chae, B. (2025). Students' trust in AI-generated explanations: Exploring cognitive vulnerabilities. *Education and Information Technologies, 30*, 9223–9244. https://doi.org/10.1007/s10639-024-13268-7

Kritkharuehart, S., Vimolsilp, U., & Photipussa, P. (2024). Development of competency-based assessment model for KRU RAK THIN scholarship recipients. *Cogent Education, 11*(1), 2373230. https://doi.org/10.1080/2331186X.2024.2373230

Kushki, A., & Nassaji, H. (2024). L2 reading assessment from a sociocultural theory perspective: The contributions of dynamic assessment. *Education Sciences, 14*(4), 342. https://doi.org/10.3390/educsci14040342

Lee, S., Choi, H., & Kim, M. (2025). Understanding student prompting behaviors with generative AI in university coursework. *British Journal of Educational Technology, 56*(1), 89–108. https://doi.org/10.1111/bjet.13492

Li, D., & Qian, J. (2024). Engaging in dialogic peer feedback in L2 writing development: A microgenetic approach. *System, 126*, 103513. https://doi.org/10.1016/j.system.2024.103513

Liu, X., Scheffel, M., & Drachsler, H. (2024). Identifying collaborative learning dynamics through chat-based analytics. *IEEE Transactions on Learning Technologies, 17*(1), 112–127. https://doi.org/10.1109/TLT.2023.3328795

Mahmud, T., Rabbi, M., Hossain, M. T., & Talukder, A. A. (2025). Portfolio assessment for developing higher order thinking skills in Bangladeshi undergraduate EFL writing classes. *Language Testing in Asia, 15*, 61. https://doi.org/10.1186/s40468-025-00404-6

Mani, Z. A. (2025). Transitioning to competency-based education in nursing: A scoping review of curriculum review and revision strategies. *BMC Nursing, 24*, 1111. https://doi.org/10.1186/s12912-025-03319-y

Mao, H., Tang, W., & Huang, L. (2024). Teaching critical AI literacy through structured critique: Reducing susceptibility to AI hallucinations. *Computers in Human Behavior, 149*, 107951. https://doi.org/10.1016/j.chb.2024.107951

McCulloch, A., Loeser, C., & Bageas, R. (2025). Does the viva matter? PhD student experiences of the oral examination and its contribution to examination outcome and researcher development. *Assessment & Evaluation in Higher Education*. Advance online publication. https://doi.org/10.1080/02602938.2025.2566652

McNamara, D. S., Allen, L. K., Crossley, S. A., Dascalu, M., & Perret, C. A. (2017). Natural language processing and learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gašević (Eds.), *Handbook of learning analytics* (pp. 93–104). Society for Learning Analytics Research. https://doi.org/10.18608/hla17.008

Min, J., Park, Y., & Cho, K. (2024). Patterns of thinking in AI-assisted academic writing: A prompt-log analysis. *Journal of Writing Research, 16*(2), 233–260. https://doi.org/10.17239/jowr-2024.16.02.03

Mori, S., Zhao, L., & Kang, J. (2024). Expertise-dependent effects of AI support in problem-based learning. *Computers & Education, 210*, 104905. https://doi.org/10.1016/j.compedu.2024.104905

Murcia, D., Jaime-Osorio, M. F., Jaramillo-Calderón, L. F., & Hoyos Pipicano, Y. A. (2025). GenAI in language teaching, learning, and assessment: Stakeholders' insights from two undergraduate language programs. *Lenguaje, 53*(1S), e20314387. https://doi.org/10.25100/lenguaje.v53i1S.14387

Naranjo, D. M., Prieto, J. R., Moltó, G., & Calatrava, A. (2019). A visual dashboard to track learning analytics for educational cloud computing. *Sensors, 19*(13), 2952. https://doi.org/10.3390/s19132952

Nedrehagen, E., Vee, T. S., Stokstad, J. M., & Orm, S. (2025). A scoping review of dialogic formative feedback practices in higher education. *Frontiers in Education, 10*, 1696703. https://doi.org/10.3389/feduc.2025.1696703

Nguyen, H., Yu, B., & Clarke, D. (2024). Metacognitive prompting with large language models: Implications for self-regulated learning. *Educational Psychology Review, 36*, 421–451. https://doi.org/10.1007/s10648-023-09728-2

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

Northern Kentucky University Center for Teaching and Learning. (2024). *Assessment strategies: Creative and multimodal assignments*.

Ochoa, X., Huang, X., & Shao, Y. (2025). Exploring the potential of generative AI to support non-experts in learning analytics practice. *Journal of Learning Analytics, 12*(1), 65–90. https://doi.org/10.18608/jla.2025.8615

Oc, Y., Gonsalves, C., & Quamina, L. T. (2025). Generative AI in higher education assessments: Examining risk and tech-savviness on students' adoption. *Journal of Marketing Education, 47*(2), 138–155. https://doi.org/10.1177/02734753241302459

Ozer, O., & Tanrıseven, I. (2016). The effect of portfolio-based writing assessment on the development of writing skills of EFL students. *International Online Journal of Educational Sciences, 8*(3), 1–10. https://doi.org/10.15345/iojes.2016.03.004

Papamitsiou, Z., & Economides, A. (2025). Multimodal learning analytics for metacognition and self-regulation: A systematic review. *Computers in Human Behavior Reports, 17*, 101435. https://doi.org/10.1016/j.chbr.2024.101435

Parker, M. J., Anderson, C., Stone, C., et al. (2025). A large language model approach to educational survey feedback analysis. *International Journal of Artificial Intelligence in Education, 35*, 444–481. https://doi.org/10.1007/s40593-024-00414-0

Patel, S., Wong, T., & Roper, K. (2025). Hallucination awareness and verification challenges in AI-mediated learning. *Journal of Educational Computing Research, 63*(1), 98–123. https://doi.org/10.1177/07356331241234567

Pereira, D. S. M., Falcão, F., Costa, L., Lunn, B. S., Pêgo, J. M., & Costa, P. (2023). Here's to the future: Conversational agents in higher education – A scoping review. *International Journal of Educational Research, 122*, 102233. https://doi.org/10.1016/j.ijer.2023.102233

Perkins, M., & Roe, J. (2023). A practical framework for ethical AI integration in assessment. *Teaching in Higher Education, 28*(9), 1234–1250.

Perkins, M., & Roe, J. (2024). GenAI, academic integrity, and the future of assessment. *International Journal for Academic Integrity, 20*(1), 1–18.

Perkins, M., Roe, J., & Furze, J. (2024). The Artificial Intelligence Assessment Scale: A practical framework for integrating generative AI into higher education assessment. *Journal of University Teaching and Learning Practice, 21*(2), 1–23.

Petingola, M., Zhang, Y., Yan, Y., & Lin, W. (2025). Integrating ethical AI tools into educational practices for enhancing academic integrity. In *Proceedings of the 7th ACM Conference on Conversational User Interfaces* (pp. 1–10). https://doi.org/10.1145/3719160.3737626

Ponomariovienė, J., Jakavonytė-Staškuvienė, D., & Tamošiūnas, F. (2025). Implementing competency-based education through the personalized monitoring of primary students' progress and assessment. *Education Sciences, 15*(2), 252. https://doi.org/10.3390/educsci15020252

Preece, D. (2025). Voices in forensics: Exploring student perceptions of oral and podcast assessments. *Journal of Learning Development in Higher Education, 41*, 1–16.

Prinsloo, P., & Slade, S. (2017). Ethics and learning analytics: Charting the (un)charted. In C. Lang, G. Siemens, A. F. Wise, & D. Gašević (Eds.), *Handbook of learning analytics* (pp. 49–57). Society for Learning Analytics Research. https://doi.org/10.18608/hla17.004

Procter, M., Lin, F., & Heller, B. (2018). Intelligent intervention by conversational agent through chatlog analysis. *Smart Learning Environments, 5*, Article 30. https://doi.org/10.1186/s40561-018-0079-5

Renzella, J., Cain, A., & Schneider, J.-G. (2022). Verifying student identity in oral assessments with Deep Speaker. *Computers and Education: Artificial Intelligence, 3*, 100044. https://doi.org/10.1016/j.caeai.2021.100044

Rodriguez, J., Kovanović, V., & Gašević, D. (2025). Conversational learning analytics: Capturing sense-making in digital dialogue. *Journal of Learning Analytics, 12*(1), 67–92. https://doi.org/10.18608/jla.2025.9015

Samei, B., Li, H., Keshtkar, F., Rus, V., & Graesser, A. C. (2014). Context-based speech act classification in intelligent tutoring systems. In S. Trausan-Matu, K. E. Boyer, M. Crosby, & K. Panourgia (Eds.), *Intelligent tutoring systems* (pp. 235–240). Springer. https://doi.org/10.1007/978-3-319-07221-0_28

Seth, S., Sharma, R., & Bhutoria, A. (2024). Conversational dynamics in AI-supported tutoring: A comparative study with human–tutor interactions. *International Journal of Artificial Intelligence in Education, 34*(3), 621–642. https://doi.org/10.1007/s40593-023-00345-7

Shin, D., Lee, H., & Kim, S. (2024). Epistemic outsourcing to AI systems: Cognitive and educational implications. *Computers and Education: Artificial Intelligence, 5*, 100233. https://doi.org/10.1016/j.caeai.2024.100233

Siemens, G., Long, P., & Gašević, D. (2023). Educators and the evolving role of data: Reimagining professional development for learning analytics and AI. *British Journal of Educational Technology, 54*(5), 1451–1470. https://doi.org/10.1111/bjet.13323

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*(10), 1510–1529. https://doi.org/10.1177/0002764213479366

Stephenson, Z., Jackson, A., & Wilkes, V. (2023). Student experiences of the 'closed-door' PhD and doctorate level viva voce: A systematic review of the literature. *Assessment*

& Evaluation in Higher Education. Advance online publication. https://doi.org/10.1080/02602938.2023.2282941

Tanner, K. D. (2012). Promoting student metacognition. *CBE—Life Sciences Education, 11*(2), 113–120. https://doi.org/10.1187/cbe.12-03-0033

Tao, Y., & Chen, G. (2023). Coding schemes and analytic indicators for dialogic teaching: A systematic review of the literature. *Learning, Culture and Social Interaction, 39*, 100702. https://doi.org/10.1016/j.lcsi.2023.100702

Tempelaar, D. T., Rienties, B., & Nguyen, Q. (2020). Towards actionable learning analytics using dispositions. *British Journal of Educational Technology, 51*(5), 1520–1539. https://doi.org/10.1111/bjet.12946

Tsai, Y.-S., Whitelock-Wainwright, A., & Gašević, D. (2021). More than figures on your laptop: (Dis)trustful implementation of learning analytics. *Journal of Learning Analytics, 8*(3), 81–100. https://doi.org/10.18608/jla.2021.7379

University of Illinois Center for Innovation in Teaching and Learning. (2021). *Multimodal assignments: Fun for students and instructors*.

Uzun, Y., Suraworachet, W., Zhou, Q., Gauthier, A., & Cukurova, M. (2025). Engagement with analytics feedback and its relationship to self-regulated learning competence and course performance. *International Journal of Educational Technology in Higher Education, 22*, 17. https://doi.org/10.1186/s41239-025-00515-3

Vargas, H., Arredondo, E., Heradio, R., & de la Torre, L. (2025). Standardizing course assessment in competency-based higher education: An experience report. *Frontiers in Education, 10*, 1579124. https://doi.org/10.3389/feduc.2025.1579124

Vázquez-Parra, J. C., Henao-Rodríguez, C., Lis-Gutiérrez, J. P., Palomino-Gámez, S., & Suárez-Brito, P. (2024). Perception of AI tool adoption and training: Initial validation using GSEM method. *Applied Computing and Informatics*.

Wang, D., & Han, H. (2021). Applying learning analytics dashboards based on process-oriented feedback to improve students' learning effectiveness. *Journal of Computer Assisted Learning, 37*(2), 487–499. https://doi.org/10.1111/jcal.12502

Wen, M., Yang, D., & Rosé, C. P. (2014). Linguistic reflections of student engagement in massive open online courses. In *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media* (pp. 525–534). AAAI. https://doi.org/10.1609/icwsm.v8i1.14512

Wheaton, M. (2024). Sociocultural learning theories for social-ecological change. *Environmental Education Research, 30*(x), 1–18. https://doi.org/10.1080/13504622.2024.2347888

Williams, A. (2025). Integrating artificial intelligence into higher education assessment. *Intersection: A Journal at the Intersection of Assessment and Learning, 6*(1), 128–154. https://doi.org/10.61669/001c.131915

Wu, X., Saraf, P. P., Lee, G., Latif, E., Liu, N., & Zhai, X. (2025). Unveiling scoring processes: Dissecting the differences between LLMs and human graders in automatic scoring. *Technology, Knowledge and Learning*, 1–16.

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education, 21*(1), 40. https://doi.org/10.1186/s41239-024-00468-z

Xu, Z., Choudhury, F. H., Ma, S., Murphrey, T. P., & Dooley, K. E. (2025). Leveraging learning analytics to model student engagement in graduate statistics: A problem-based learning approach in agricultural education. *Behavioral Sciences, 15*(10), 1360. https://doi.org/10.3390/bs15101360

Zhang, X., Wong, M., & Li, J. (2024). Students' elaborative prompting and critical reasoning in AI-supported academic tasks. *Computers & Education, 209*, 104826. https://doi.org/10.1016/j.compedu.2024.104826

Zou, H., Luo, L., & Li, X. (2024). Large language model-enhanced writing analytics: Detecting cognitive moves in student-AI conversation. *Assessment & Evaluation in Higher Education, 49*(3), 311–329.

Chapter 2

Digital Empowerment for Students with Visual Impairments: Accessible Design, Assistive Innovation, and Inclusive Higher Education

Sahar Abu Shokhedim[1]
and Jamil Itmazi[2,*]
[1]An-Najah National University, Nablus, Palestine
[2]Palestine Ahliya University, Bethlehem, Palestine

Abstract

Visual impairments remain among the most significant barriers to equitable access in higher education, particularly as institutions worldwide have rapidly transitioned from face-to-face instruction to digital learning environments. While digital platforms offer flexibility, personalization, and scalability, they may also intensify exclusion when accessibility is treated as an optional add-on rather than a foundational design principle. This chapter presents a comprehensive examination of the digital empowerment of students with visual impairments by integrating theoretical, technological, pedagogical, and policy-related perspectives. Guided by key learning frameworks, including Universal Design for Learning (UDL) and Diffusion of Innovations Theory, the chapter explores innovative approaches to addressing diverse visual needs through accessible design, multisensory learning, institutional readiness, and the strategic deployment of assistive and emerging technologies.

The chapter examines four major categories of visual impairment—night blindness, color vision deficiency, low vision, and total blindness—

* Corresponding Author's Email: j.itmazi@paluniv.edu.ps

In: Progress in Education. Volume 91
Editor: Roberta V. Nata
ISBN: 979-8-90134-375-3

highlighting their functional characteristics and educational implications. It further analyzes the multiple layers of inaccessibility encountered by visually impaired students, including technological, pedagogical, institutional, financial, and socio-cultural barriers, Subsequently, the chapter reviews a wide spectrum of enabling technologies, ranging from conventional assistive tools to emerging solutions such as artificial intelligence, wearable devices, haptic interfaces, extended reality, and brain-computer interfaces. Comparative case studies from Palestine, Ethiopia, South Africa, Canada, and Finland illustrate how different educational systems address accessibility through localized solutions, training initiatives, governance structures, and policy integration. These findings are synthesized into an integrative model of digital empowerment encompassing six interrelated dimensions. policy integration, infrastructure, agency, autonomy, access, and capacity building. The chapter concludes by outlining future pathways, including haptic learning ecosystems, re-engineered tactile e-books, blockchain-enabled credentials, and AI-driven tutoring systems that hold potential for transforming digital education into a more inclusive and equitable ecosystem. Overall, the chapter argues that digital empowerment extends beyond tools and technologies, reflecting instead the extent to which higher education institutions demonstrate a sustained commitment to inclusive, equitable, and accessible teaching and learning.

Keywords: digital empowerment, visual impairments, assistive technology, inclusive education, Universal Design for Learning (UDL), accessibility, artificial intelligence

1. Introduction

1.1. Preface

Vision plays a central role in how individuals perceive, learn from, and interact with their environment. It enables the interpretation of spatial relationships, written text, symbols, facial expressions, and navigation within both physical and digital space. In higher education, teaching and learning processes are heavily dependent on visual information. Instruction typically relies on visual materials such as lecture slides, diagrams, tables, laboratory demonstrations, virtual interfaces, and images, often assuming that learners can efficiently process visual content alongside spoken explanations.

Such visual expectations pose significant challenges for students with mild, moderate and severe visual impairments. Even minor visual limitations may substantially increase the cognitive and physical effort required for comprehension , participation, and completion of academic tasks that rely on efficient access to visual information.

According to the World Health Organization (WHO, 2023), more than 2.2 billion people worldwide experience some degree of visual impairment, making it one of the most prevalent forms of disability globally. The rapid transition of higher education toward digital learning environments has generated both new opportunities and substantial challenges. Digital technologies may enhance flexibility and personalization, while simultaneously integrating assistive features such as screen readers, magnification tools, tactile interfaces, audio feedback and AI-based recognition systems.

However, poorly designed digital platforms, however—particularly those lacking semantic structure or descriptive alternatives—may exclude students with visual impairments even more severely than traditional instructional materials. As noted by UNESCO (2023), digital inequalities have intensified in a number of countries, largely due to uneven access to resources and the limited capacity of disability support units in many institutions.

1.2. Digital Empowerment as a Concept

Digital empowerment extends beyond mere access to technological tools. It represents a transformative process through which students with visual impairments develop the capacity to engage meaningfully, independently, and confidently within digital learning environments. It includes:

- Accessible design of digital content and platforms,
- Availability and accessibility of assistive devices,
- institutional policies, training and supports in place,
- Opportunities to develop learning tools together, or personalize them,
- Cultural acceptance and inclusive pedagogies,
- Reliable infrastructure and maintenance.

Digital empowerment could be seen as a multiaxial program, encompassing technology, pedagogical models, policy frameworks and user focus designs. This view is in line with those worldwide, who are advocating for inclusive education also in light of "SDG 4"and the UN convention on the rights of persons with disabilities.

Figure 1 illustrates the inputs, processes, outputs, and feedback mechanisms of the digital empowerment cycle for students with visual impairments.

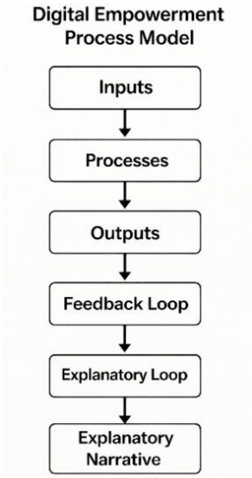

Figure 1. Digital Empowerment Process Model.

This chapter adopts a holistic perspective on digital empowerment for students with visual impairments in higher education. It integrates relevant theoretical frameworks with detailed explanations of the types and patterns of visual impairment, as well as an expanded discussion of their functional and educational implications. The chapter aims to provide a robust conceptual foundation for the subsequent practical and technical discussions. Methodologically, the chapter employs an analytical synthesis approach that comprises three core elements: a theoretical review, a set of comparative global case studies, and the development of a conceptual model of digital empowerment.

2. Theoretical Foundations of Digital Empowerment

Understanding digital empowerment requires engagement with multiple theoretical perspectives that explain learning processes, technology adoption, and inclusive educational practices, collectively offering a conceptual lens for examining accessibility, innovation, and participation in higher education.

2.1. Universal Design for Learning (UDL)

Universal Design for Learning (UDL) represents one of the most influential frameworks guiding inclusive educational practice. As articulated by Meyer, Rose, and Gordon (2014), UDL advocates for learning environments that are accessible to the widest possible range of learners from the outset the outset rather than relying on retrofitted or individualized accommodation. The UDL framework is structured around three core principles:

1. Multiple Means of Presentation: Providing multiple ways of presenting information. For learners who are blind or visually impaired, this may include:
   - Audio descriptions,
   - Tactile graphics,
   - Text-to-speech formats,
   - Enlarged print,
   - High-contrast visuals,
   - Non-visual alternatives to color-based information.
2. Multiple Means of Action and Expression: Allowing students to demonstrate learning outcomes through flexible and accessible formats, including:
   - Audio recordings,
   - Screen-reader–compatible digital submissions,
   - Oral presentations,
   - Tactile or physical models.
3. Multiple Means of Engagement: Supporting learner motivation and autonomy by offering:
   - Choices in how content is accessed and explored,
   - Personalization options,
   - Technologies responsive to individual learning needs.

UDL shifts the focus from “fixing” the learner to adapting the learning environment itself. It explicitly recognizes that learning barriers arise from design limitations rather than from learners’ disabilities.

2.2. Diffusion of Innovation (DOI)

Rogers' (2003) Diffusion of Innovation (DOI) theory explains how new technologies are adopted within social systems and organizations. The theory is particularly relevant to assistive technologies and accessibility initiatives, as it helps explain variations in adoption across individuals and institutions.

According to DOI, adoption is influenced by several key factors:

- Relative advantage: Does the innovation provide clear benefits over existing practices?
- Compatibility: *Is it aligned with institutional values and teaching norms*?
- Complexity: Is it easy to understand and implement?
- Trialability: Can it be tested on a limited scale before full adoption?
- Observability: Are its benefits visible to others?

While students with visual impairments often adopt assistive technologies rapidly, institutions typically demonstrate slower uptake, particularly in the presence of policy gaps, insufficient training, or funding constraints. This institutional inertia has been documented in several higher education contexts (Khumalo & Naidoo, 2024).

2.3. Inclusive Education and the Social Model of Disability

The social model of disability offers a holistic perspective which argues that disability emerges from the interaction between individual functional limitations and inaccessible environment. Applied to visual impairment, this model shifts attention away from medical diagnoses toward environmental and systemic barriers, such as:

- Inaccessible learning management systems,
- Poorly designed digital instructional materials,
- Limited institutional training and awareness,
- Inadequate faculty preparation for inclusive teaching.

From this perspective, digital empowerment requires sustained commitment. Rather than relying on individualized adjustments, universities must critically examine and enhance their digital systems, policies, and instructional practices.

3. Visual Impairments: Definitions, Features, and Educational Relevance

Visual impairment encompasses a broad spectrum of functional limitations that affect how individuals perceive, process, and interpret visual information, rather than representing a single, uniform condition. Understanding the distinct characteristics of each condition is essential for designing effective digital and educational adaptations.

3.1. Night Blindness (Nyctalopia)

Night blindness (nyctalopia) refers to difficulty seeing in low-light conditions or adapting quickly between bright and dim environments (Cleveland Clinic, 2024). Common causes include:

Rod photoreceptor dysfunction:
- Retinitis pigmentosa,
- Cataracts,
- Vitamin A deficiency,
- Glaucoma or retinal diseases,

Functional characteristics include difficulties with:
- Navigating dimly lit environments,
- Participating in evening or low-light classes,
- Reading screens in reduced lighting,
- Adjusting to sudden changes in brightness.

Educational Implications require universities to provide:
- Adequate and adjustable lighting in learning spaces,
- Interfaces with stable brightness controls,
- High-contrast digital materials,
- Additional time for visual adjustment between environments.

In digital environments, high-contrast designs and customizable brightness settings can substantially improve accessibility.

3.2. Color Vision Deficiency (Color Blindness)

Color vision deficiency refers to a reduced ability to distinguish between certain colors. The primary forms include:

- Red–green deficiency,
- Blue–yellow deficiency,
- Complete monochromacy (rare).

Approximately 0.5% of women and 8% of men experience some form of color-vision deficiency. (National Eye Institute, 2023; Color Blindness Awareness, n.d.).

Functional Characteristics include challenges in:

- Interpreting color-coded charts and graphs,
- Understanding laboratory indicators,
- Using color-dependent interface buttons,
- Decoding diagrams that rely heavily on color cues.

Educational Implications include the need to:

- Use labels, patterns, and shapes in addition to color,
- Avoid problematic color combinations,
- Ensure compliance with WCAG contrast requirements,
- Design dashboards and visualizations with alternative non-color indicators.

Although often overlooked, color vision deficiency can have a substantial impact on access within digital learning environments.

3.3. Partial Sight (Low Vision)

Low vision refers to long-term visual impairment that cannot be fully corrected through glasses, contact lenses, or surgical interventions. It may affect visual acuity, peripheral vision, contrast sensitivity or depth perception (American Foundation for the Blind, n.d.).

Functional characteristics commonly include:

- Blurred or distorted vision,
- Reduced visual field,
- Difficulty reading standard-sized text,
- Challenges recognizing distant objects,
- Increased visual fatigue and eye strain.

Educational implications suggest that students with low vision benefit from:

- Magnification software and tools,
- Screen readers with speech output,
- Reflowable digital text formats,
- High-resolution displays,
- Large-print instructional materials,
- High-contrast interface themes.

When implemented effectively, font scaling and zoom features in modern digital platforms can provide substantial support for learners with low vision.

3.4. Total Blindness

Total blindness refers to the complete absence of light perception (MedlinePlus, 2024). Students who are totally blind rely exclusively on non-visual forms of access.

Key functional characteristics include:

- No visual access to digital interfaces,
- Dependence on auditory and tactile feedback,
- Extensive use of screen-reader software such as JAWS or NVDA,
- Reliance on keyboard navigation and structured HTML content.

Educational implications require universities to ensure:

- Full screen-reader compatibility,
- Meaningful and descriptive alternative text,

- Clear and logical content structure,
- Availability of tactile graphics,
- Audio-described multimedia content,
- accessible digital assessments.

Digital materials that lack semantic structure or description content can effectively exclude blind learners from meaningful participation.

3.5. Comparative Educational Implications

Each category of visual impairment generates distinct educational needs. Accordingly, accommodations should be based on functional impact rather than solely on medical diagnosis.

Condition	Functional Limitation	Key Accommodations
Night Blindness	Low-light navigation	High contrast, stable brightness, lighting control
Color-Vision Deficiency	Color discrimination	Icons, labels, non-color cues
Low Vision	Low contrast, small text	Zoom, magnifiers, reflowable text
Total Blindness	No visual access	Screen readers, tactile graphics, audio

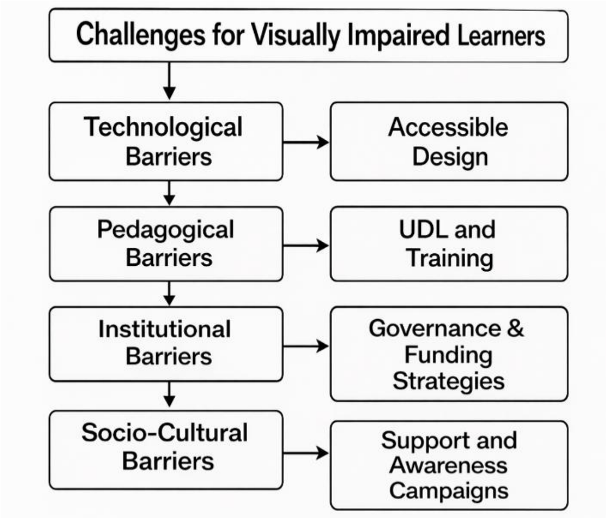

Figure 2. Challenges and Corresponding Solutions for Visually Impaired Learners.

4. Obstacles to Digital Inclusion for Students with Visual Impairment

Students with visual impairments encounter multiple interrelated challenges that affect their participation and engagement in digital learning environments. These challenges span technological, pedagogical, institutional, financial, and socio-cultural dimensions. Identifying these barriers is essential for developing effective and context-sensitive inclusion strategies.

Figure 2 summarizes key barriers (technological, pedagogical, institutional, socio-cultural), and their corresponding solutions.

4.1. Technological Barriers

Technological barriers represent one of the most pervasive challenges faced by students with visual impairments.

4.1.1. Inaccessible Learning Management Systems (LMS)

Many learning management systems—particularly legacy institutional platforms—were developed without adequate accessibility considerations, resulting in:

- Unlabeled Buttons and menus,
- Images or icons lacking alternative text,
- PDF documents that are not OCR- compatible,
- Drag-and-drop activities inaccessible via keyboard navigation,
- Multimedia content without captions or audio descriptions,
- Tables read in illogical order by screen readers.

These limitations often force students to rely on ad hoc solutions or assistance from others, thereby reducing learner autonomy.

4.1.2. Weak Compatibility with Assistive Technologies

For optimal functionality, screen readers software, such as NVDA or JAWS requires clearly structured digital content, including appropriate heading hierarchies and accurately labeled interface elements on the page.

However, many university websites fail to meet these basic requirements. As noted by Petrie and Bevan (2009), even minor coding errors, such as empty headings or improper nesting, can create substantial access barriers.

4.1.3. Digital Environments Designed Primarily for Sighted Users

Many digital platforms are designed primarily for sighted users and rely heavily on visual features, including:

- Color-coded indicators,
- Small or unlabeled icons,
- Low contrast design elements,
- Hover-dependent interactions,
- image-based navigation systems.

Such design assumptions marginalize students who depend on non-visual or multimodal interaction methods.

4.1.4. Lack of Accessible Formats

A considerable proportion of academic content remains available only in formats inaccessible to screen readers, including:

- Scanned PDFs documents,
- Images containing text without descriptions,
- Infographics lacking textual alternatives,
- Charts presented without accompanying data tables.

As a result, visually impaired students frequently depend on others to convert materials into accessible formats, who then have to convert the materials into an accessible format, a process that is time-consuming, inequitable, and detrimental to independent learning.

4.2. Pedagogical Barriers

Instructional practices play a critical role in shaping the educational experiences of students with visual impairments.

4.2.1. Overreliance on Visual Instruction

Many instructors rely heavily on visual instructional materials to communicate such as slides, charts, diagrams, and videos, without providing sufficient verbal or tactile explanations.

While such materials may support sighted learners, they become significant barriers when alternative formats are not provided (Kelly & Smith, 2021).

4.2.2. Lack of Faculty Training

In many universities, instructors receive limited or no formal training in:

- Accessible course design,
- Writing effective alternative text,
- Creating screen-reader– compatible documents,
- Applying UDL principles in instructional design.

When instructors lack awareness of accessibility standards, they may unintentionally develop materials that hinder access for visually impaired learners.

4.2.3. Assessment Barriers

Assessments practices frequently include:

- Visually dependent interpretation tasks,
- Graph-reading questions without textual descriptions,
- Color-coded instructions,
- Timed online tests incompatible with screen readers.

These assessment formats may disadvantage students with visual impairments, even when they demonstrate full mastery of the academic content.

4.3. Institutional Barriers

Institutional structures and governance significantly influence whether digital empowerment can be effectively realized.

4.3.1. Absence of Accessibility Policies

In many universities, particularly in developing contexts, formal accessibility policies are either weakly articulated or entirely absent. This lack of policy clarity results in:

- Unclear institutional requirements,
- Limited accountability mechanisms,
- Inconsistent accessibility practices across departments.

4.3.2. Limited Disability Support Services

Although some institutions maintain disability support units, these services are frequently:

- Understaffed,
- Underfunded,
- Limited in expertise with assistive technologies,
- Focused primarily on examination accommodations rather than digital access.

The National Center for College Students with Disabilities (NCCSD, 2021) emphasizes the importance of well-resourced and institutionally integrated accessibility offices.

4.3.3. Technological Infrastructure Gaps

Students frequently report several infrastructure-related challenges, including:

- Unreliable campus Wi-Fi,
- Slow laboratory computers,
- Outdated software,
- Restrictions on installing screen readers on university devices.

Such limitations significantly reduce the effectiveness of available assistive technologies, even when students are adequately trained to use them.

4.4. Financial Barriers

Financial constraints represent one of the most significant barriers to digital inclusion largely because many assistive technologies involve substantial costs.

4.4.1. High Cost of Assistive Software

Several widely used assistive tools require high initial investments, including:

- JAWS screen reader (over USD 1,000),
- ZoomText magnification software (several hundred USD),
- Refreshable Braille displays (often exceeding USD 3,000).

These costs are prohibitive for many students, particularly in low- and middle-income contexts (Hartley, 2018).

4.4.2. Maintenance and Replacement Costs

Even when institutions invest in purchasing assistive technologies, they often fail to:

- Renew software licenses,
- Replace malfunctioning devices,
- Provide ongoing training and support,
- Ensure regular technical maintenance.

As a result, malfunctioning or outdated devices accumulate, undermining the long-term sustainability of digital empowerment initiatives.

4.5. Social and Cultural Barriers

Social attitudes and cultural perceptions strongly shape the learning experiences of students with visual impairments.

4.5.1. Stigma and Misconceptions

Students frequently encounter assumptions that they:

- Are unable to achieve academic excellence,

- Require constant support,
- Cannot independently use digital technologies.

Such misconceptions undermine students' confidence, participation, and sense of belonging (Daniels & Scott, 2020).

4.5.2. Reduced Opportunities for Collaboration

Collaborative learning and peer interaction may exclude visually impaired students when:

- Materials are shared in inaccessible formats,
- Group communication relies on visuals cues,
- Digital collaboration platforms lack screen-reader compatibility.

4.5.3. Social Isolation

Feelings of social isolation may intensify when:

- Accessibility is not prioritized,
- Students hesitate to request accommodations,
- Faculty and peers lack awareness of accessibility needs.

Consequently, social inclusion represents a critical dimension of sustainable digital empowerment.

5. Web Accessibility for Users with Visual Impairments

The web constitutes a primary medium for learning, communication, and academic engagement in higher education. The accessibility of web-based environments determines whether students with visual impairments can effectively navigate learning management systems, access online content, submit assignments, and participate in digital activities. International accessibility efforts are guided primarily by the Web Accessibility Initiative (WAI) and the Web Content Accessibility Guidelines (WCAG) 2.1.

5.1. Total Blindness and Screen-Reader Interaction

Students who are totally blind typically rely on screen readers such as NVDA, JAWS, or VoiceOver to access digital content. These tools convert on-screen text into speech or Braille output and function effectively only when websites comply with established accessibility standards.

5.1.1. Common Barriers

Common barriers include:

- Missing inadequately written alternative text,
- Images conveying meaning without textual explanation,
- Decorative images incorrectly marked as informative,
- Improper use of ARIA roles,
- Poorly structured headings hierarchies that hinder navigation.

Lazar et al. (2015) note that blind users frequently encounter digital content that lacks coherent structure, resulting in significant navigation and comprehension difficulties.

5.1.2. Accessibility Requirements

Effective accessibility requires:

- Descriptive alternative text that communicates purpose,
- Logical heading hierarchy (H1–H6),
- Clearly labeled and searchable form fields,
- Skip-navigation links to bypass repeated content,
- Avoidance of image-only buttons for interactive controls.

Together, these elements enable independent navigation and contribute directly to digital empowerment.

5.2. Low Vision and Magnification Compatibility

Students with low vision commonly depend on:

- Browser-based zoom features,
- Screen magnification tools,
- High-contrast display modes,
- Customizable text size and spacing settings.

5.2.1. Common Challenges
Frequent challenges include:

- Layout distortion during magnification,
- Overlapping or truncated interface elements,
- Text that fails to scale proportionally,
- Low-contrast color combinations.

Petrie and Bevan (2009) emphasize that magnification often exacerbates design flaws, thereby reducing overall accessibility.

5.2.2. Accessibility Requirements
Accessible design for low-vision users requires:

- Responsive layouts that support magnification,
- High-contrast text and background combinations,
- Clear typography and adequate spacing,
- Reflowable text that adapts to enlarged display settings.

5.3. Night Blindness (Nyctalopia)

Night blindness affects users' ability to perceive content in dim or low-contrast digital environments.

5.3.1. Web Challenges Include
Key challenges include:

- Dark-mode interfaces with insufficient contrast,
- Flickering or animated backgrounds,
- Distracting gradients and visual noise,
- Color schemes that impair visibility.

5.3.2. Accessibility Strategies
Effective accessibility strategies include:

- Providing both light and dark display modes,
- Maintaining stable high-contrast design,
- Avoiding flashing or rapidly changing elements,
- Ensuring visual consistency across pages.

5.4. Color Vision Deficiency

Information should never rely on color as the sole means of communication.

5.4.1. Common Barriers
Common barriers include:

- Red/green indicators for status or correctness,
- Color-dependent charts and graphs,
- Traffic-light–style interfaces,
- Low-contrast color pairings.

Birch (2012) and Color Blindness Awareness emphasize that color-only communication renders critical information inaccessible to many users.

5.4.2. Accessibility Requirements
Accessible design for color vision deficiency requires:

- Icons, labels, or patterns alongside color,
- Color-safe and perceptually distinct palettes,
- Strong contrast ratios,
- Alternative non-color indicators for status and feedback.

6. Assistive Technologies and Their Empowering Role

Assistive technologies play a central role in converting visual information into accessible formats, thereby enhancing independence, confidence, and academic self-efficacy among students with visual impairments.

Assistive technologies for visually impaired learners can be broadly classified into six major categories as illustrated in Figure 3 and are used by visually impaired learners.

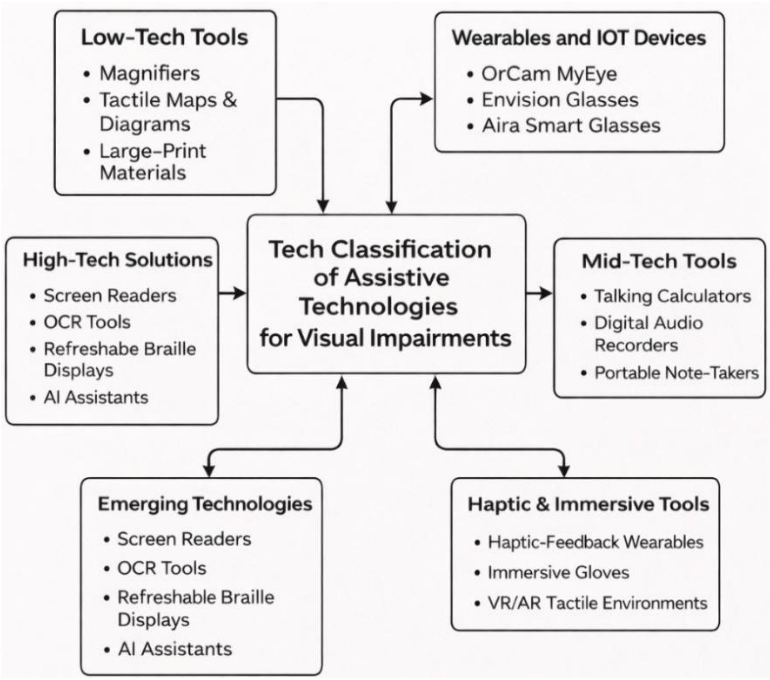

Figure 3. A Classification of Assistive Technologies for Visual Impairments.

6.1. Low-Tech, Mid-Tech, and High-Tech Solutions

Figure 4 presents a functional classification of assistive technologies according to their level of technological complexity.

1. *Low-Tech:* Low-tech solutions include simple, non-digital tools such as
   - Magnifiers,
   - Large-print books and materials,
   - Tactile diagrams and raised-line graphics.
2. *Mid-Tech:* Mid-tech solutions include devices that provide audio-based support, such as

- Talking calculators,
- Digital audio recorders,
- Portable note-taking devices.

3. *High-Tech:* High-tech solutions involve advanced digital and AI-enabled tools, including
   - Screen readers such as JAWS and NVDA,
   - OCR and text-recognition applications (e.g., Seeing AI),
   - Refreshable Braille displays (e.g., BrailleNote Touch+),
   - AI-based virtual assistants.

These technologies are most effective when integrated into students' daily academic routines (Al-Azawei & Serenelli, 2021).

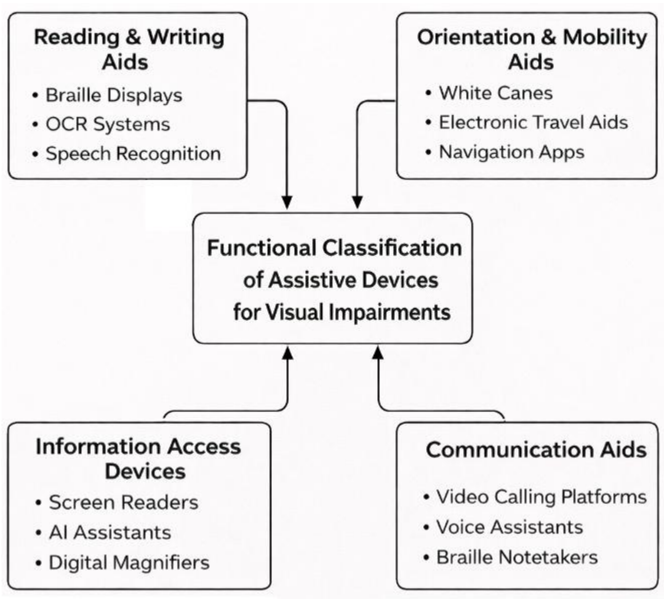

Figure 4. Functional Classification of Assistive Devices for Visual Impairments.

6.2. Wearable Technologies

Wearable assistive technologies increasingly combine cameras, artificial intelligence, and audio output to support real-time access to visual information.

6.2.1. OrCam MyEye

OrCam MyEye is a wearable device capable of reading printed text aloud, recognizing faces, and identifying everyday objects.

6.2.2. Envision Glasses

Envision Glasses provide real-time text recognition, object identification, and environmental description through AI-driven computer vision.

6.2.3. Aira Smart Glasses

Aira Smart Glasses connect users with trained human agents who provide live visual interpretation and navigation assistance (Pune Blind School, 2024).

By enhancing situational awareness and mobility, these wearable technologies promote greater confidently and more confident participation in academic and social activities.

6.3. Haptic and Tactile Technologies

Haptic and tactile technologies offer non-visual alternatives for accessing spatial, auditory, and structural information.

Examples include:

- NeoSensory Buzz, which converts sound frequencies into vibrations tactile,
- HaptX Gloves, which provide immersive spatial and tactile simulations.

Although often underutilized in higher education, haptic technologies have demonstrated particular value in STEM education (Kouroupetroglou et al., 2023).

7. Multisensory Learning and UDL-Based Strategies

Multisensory learning integrates auditory, tactile, and kinesthetic modalities to provide alternative pathways for information processing, particularly benefiting learners with severe visual impairments. By reducing reliance on visual input, this approach enhances comprehension, engagement, and knowledge retention within inclusive higher education contexts.

7.1. Tactile and Auditory Enhancements

The use of 3D models, tactile graphics, sonified data, audio navigation systems, and haptic feedback, provides viable alternatives to visual input for learners with visual impairments. Engagement with these formats can lead to deeper conceptual understanding while also reducing cognitive load during learning tasks. Tactile graphics are particularly effective in disciplines requiring spatial reasoning, such as mathematics, engineering, and the natural sciences. Complementing these tools, audio-based navigation systems and screen-reader support enable students to independently explore complex digital environments, thereby enhancing accessibility and academic engagement.

7.2. Significance in STEM and Digital Environments

STEM disciplines traditionally rely heavily on visual representations. Multisensory strategies provide alternative pathways for learning by enabling students to comprehend spatial structures, explore scientific models, analyze data through non-visual representations, and engage meaningfully in simulations. For example, sonification tools translate graphs and numerical patterns into auditory cues, allowing blind students to interpret trends, patterns, and variations independently. Similarly, haptic simulations represent geometric forms or molecular structures through touch-based interaction. These approaches not only enhance access but also promote deeper conceptual learning for a broad range of learners.

7.3. An Alignment with the Universal Design for Learning

Multisensory learning aligns closely with the core principles of Universal Design for Learning (UDL). It operationalizes multiple means of representation by presenting information through varied and complementary sensory channels, thereby reducing dependence on a single mode of access. It further supports multiple means of engagement by enhancing learner motivation, autonomy, and sustained participation. When implemented systematically across curricula, multisensory strategies contribute to inclusive learning environments that benefit students with and without disabilities alike.

8. The Emerging Technologies in Accessibility for Visually Impaired Students

Recent advances in digital technologies have created unprecedented opportunities for accessibility and support for students with visual impairments. While traditional assistive tools—such as magnifiers and screen readers—remain essential, developments in artificial intelligence, computer vision, wearable technologies, extended reality, and brain–computer interfaces have significantly expanded the scope of accessible learning, particularly with respect to independent learning and digital participation. These technologies do not merely provide assistance; rather, they actively transform the educational landscape by redefining the capabilities and learning experiences of students with visual impairments.

8.1. Artificial Intelligence (AI) and Computer Vision

Artificial intelligence is increasingly central to mitigating vision-related limitations in educational contexts. AI-driven recognition tools can interpret visual information in real time, thereby rendering previously inaccessible materials accessible to students with visual impairments.

8.1.1. Scene and Object Recognition

Applications such as Seeing AI, Google Lookout, and Be My Eyes (AI Mode) employ machine-learning techniques to:

- Describe scenes,
- Recognize faces,
- Identify objects,
- Read printed and handwritten text.

These capabilities enable students to independently access visual information within classrooms, libraries, and other learning environments. (Pfeiffer et al., 2022).

8.1.2. Text Recognition and Digital Reading

Advances in machine learning have significantly improved the speed and accuracy of optical character recognition (OCR) technologies. Tools such as Seeing AI and KNFB Reader can:

- Convert printed text into speech,
- Navigate multi-column page documents,
- Interpret complex layouts,
- Process PDF files that were previously inaccessible.

These developments expand access to instructional materials and strengthen digital reading skills among visually impaired learners.

8.2. Smart Assistants and Conversational AI

Conversational AI systems—such as ChatGPT, Google Assistant, Apple Siri, and Amazon Alexa—offer substantial support for visually impaired students.

8.2.1. Academic Support

Smart assistants can support a range of academic tasks, including:

- Summarizing academic texts,
- Explaining complex concepts,
- Generating illustrative examples,
- Converting text into speech.

By reducing dependence on human intermediaries, these tools support self-regulated learning is supported (Selwyn, 2020).

8.2.2. Organizational Support

AI assistants further support organizational and navigational tasks, such as:

- Scheduling and time management,
- Task reminders,
- Non-visual navigating digital interfaces,
- Web searching through voice-based interaction.

These functions assist students in managing daily academic demands and enhance overall independence.

8.3. Wearable Technologies: The Convergence of AI and Mobility

Among emerging accessibility solutions, wearable technologies represent some of the most transformative innovations.

8.3.1. OrCam MyEye

OrCam MyEye is a compact, eyeglass-mounted camera that:

- Reads text aloud,
- Identifies products,
- Recognizes faces,
- Provides discreet audio feedback.

8.3.2. Envision Glasses

Envision Glasses offer:

- Real-time text reading,
- Object recognition,
- Character and handwriting interpretation,
- Environmental descriptions through audio output (Envision, 2024).

8.3.3. Aira Smart Glasses

Aira Smart Glasses connect users with trained human agents who provide:

- Live visual interpretation,
- Assistance with navigation,
- Task-specific guidance (Pune Blind School, 2024).

Collectively, these wearable technologies enhance mobility, support classroom participation, and facilitate independent problem-solving ability.

8.4. Haptic Technologies and Tactile Interaction

Haptic interfaces deliver tactile stimuli that enable students to perceive spatial and structural information through touch.

8.4.1. Vibro-Feedback Devices

Devices such as NeoSensory Buzz convert sound frequencies into tactile vibrations, thereby enhancing environmental awareness.

8.4.2. Immersive Haptic Gloves

Immersive devices such as HaptX Gloves enable users to perceive textures, movement, and spatial structures. These systems support:

- Manipulation of virtual objects,
- Deeper engagement with STEM concepts,
- Rich multisensory learning experiences (Kouroupetroglou et al., 2023).

8.5. Extended Reality (XR) and Accessible Immersion

Virtual Reality (VR) and Augmented Reality (AR) need not rely exclusively on visual input. When integrated with tactile and auditory feedback, these technologies can offer novel educational experiences for blind and visually impaired learners. Examples include:

- Audio-based virtual environments,
- Tactile overlays for three-dimensional exploration,
- Real-time audio description with haptic guidance,
- Sound-based AR navigation systems.

UNESCO (2024) highlights initiatives such as the "Touching the Invisible" project which enables learners to explore virtual environments through touch and sound.

8.6. Brain–Computer Interfaces (BCIs)

A brain–computer interface (BCI) is a system that captures neural signals, processes them, and translates them into commands for external devices or digital systems. In educational contexts, potential applications of BCIs include:

- Navigating digital environments through neural input,
- Controlling devices without physical interaction,
- Adaptive learning systems responsive to cognitive states.

Although still largely experimental, these systems may offer new ways for learners with visual impairments to interact with digital environments (Pfeiffer et al., 2022).

9. Global Case Studies: Comparative Insights from Diverse Contexts

Comparative case studies from the Middle East (Palestine), Africa, Europe, and North America provide valuable insights into how digital empowerment operationalized across diverse educational systems and socio-economic contexts. Together, these cases reveal shared challenges as well as context-specific strategies for promoting accessibility and inclusion in higher education.

9.1. Palestine: Localized Assistive Technology and Arabic-Supported Tools

Abu Shokhedim et al. (2025) demonstrated that the adoption of localized assistive technologies in Palestinian universities resulted in:

- Improved grade point averages,
- Increased digital independence,
- Increased classroom participation,
- Reduced anxiety in digital learning environments.

The study highlights the critical importance of Arabic-language interfaces and locally grounded technical support, features that remain largely underrepresented in global assistive technology solutions.

9.2. Ethiopia: Structured Digital Empowerment Training

Tadele and Mengistu (2024) developed and evaluated a structured Digital Empowerment Training (DET) model. The implementation of the model led to:

- Significant gains in digital literacy,
- Increased confidence in using assistive technologies,
- Increased employability readiness,
- Greater engagement in independent learning.

The findings indicate that structured training programs, rather than access to technology alone, were the primary drivers of these positive outcomes.

9.3. South Africa: Governance and Sustainability Challenges

Khumalo and Naidoo (2024) observed that institutional adoption of assistive technologies was strongly influenced by:

- Effective governance structures,
- Sustainable funding mechanisms,
- Ongoing staff training,
- Dedicated maintenance and support units.

In the absence of these institutional conditions, even advanced assistive technologies risk becoming underutilized or abandoned over time.

9.4. Canada and Finland: Policy-Driven Accessibility Ecosystems

According to UNESCO (2023, 2024), Canada and Finland represent leading examples of countries that have effectively integrated digital accessibility in national education strategies.

Key Outcomes include:

- A 40% increase in the use of accessible e-libraries,
- Systematic integration of UDL principles into curricula,
- Government-funded maintenance of assistive technologies,
- National standards enhancing institutional accountability.

These cases demonstrate that sustainable accessibility improvements depend primarily on comprehensive systemic measures. rather than isolated or short-term initiatives.

Refer to Figure 5.

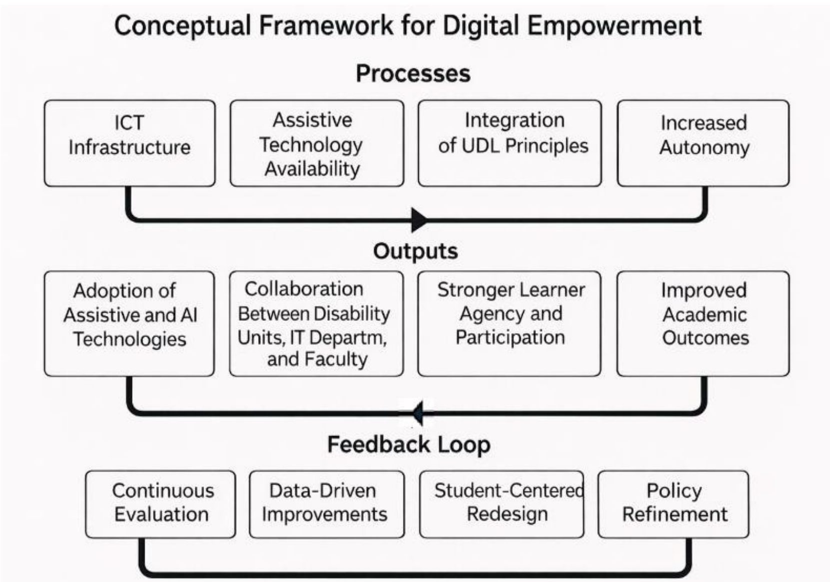

Figure 5. Conceptual Framework for Digital Empowerment (Process–Output–Feedback Model).

9.5. Research Gaps Identified Across Contexts

The comparative analysis of these case studies highlights several persistent research gaps that warrant further scholarly investigation.

Key gaps include:

- The absence of integrated, multidimensional empowerment models that combine accessibility, pedagogy, and technology within a single framework.
- Limited empirical evidence from Arab countries and socio-economically disadvantaged regions, particularly regarding localized tools and institutional readiness.
- Insufficient research on the long-term impact of emerging technologies, including AI-driven tutors, haptic devices, and extended reality environments.
- Underrepresentation of learner autonomy, agency, and psychological empowerment as outcome measures.
- Limited examination of governance, sustainability, and cost-effectiveness in assistive technology implementation.
- A lack of validated instruments for measuring digital empowerment as a multidimensional construct.

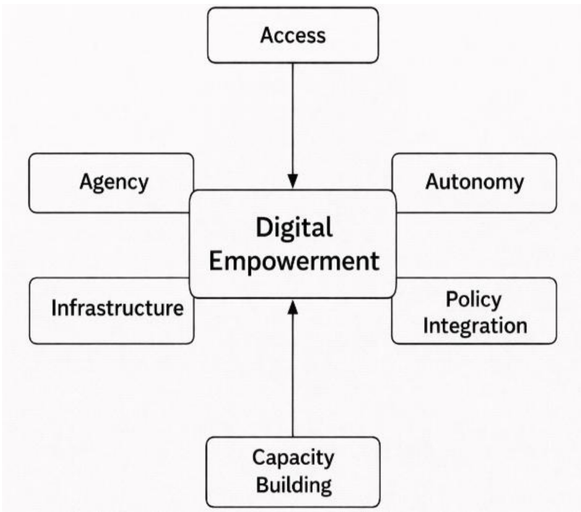

Figure 6. Six-Dimension Conceptual Framework for Digital Empowerment.

10. Digital Empowerment: A Six-Dimension Framework

Drawing on the reviewed case studies, theoretical foundations, and technological developments, this chapter proposes a comprehensive framework for digital empowerment in higher education. The framework integrates six interrelated dimensions that collectively support the digital empowerment of learners with visual impairments. Figure 6 illustrates the conceptual relationships among these dimension.

10.1. Dimension 1 – Access

Key elements of the access dimension include:

- Accessible digital learning materials,
- Full screen-reader compatibility,
- Multilingual and multimodal user interfaces,
- Reliable and inclusive ICT infrastructure.

Expected outcomes of enhanced access include:

- Equitable access to instructional content,
- Reduction of structural digital barriers,
- Inclusive participation across learning activities.

10.2. Dimension 2 – Autonomy

The autonomy dimension emphasizes:

- Personalized AI-based learning assistants,
- Tactile and multisensory learning tools,
- Real-time OCR and AI recognition systems,
- Technologies supporting self-paced learning.

Expected outcomes include:

- Increased learner independence,

- Enhanced confidence in digital environments,
- Reduced reliance on external assistance.

10.3. Dimension 3 – Agency

The agency dimension focuses on:

- Co-design of digital tools with visually impaired students,
- Participatory feedback mechanisms,
- Student-led innovation and accessibility initiatives.

Expected outcomes include:

- Active learner participation in decision-making,
- Stronger alignment between technologies and user needs,
- Greater student influence over learning experiences.

10.4. Dimension 4 – Infrastructure

Infrastructure-related elements include:

- Dedicated maintenance units for assistive technologies,
- Regular updates of hardware and software,
- Accessible campus digital ecosystems,
- Stable internet connectivity and device availability.

Expected outcomes include:

- Sustained functionality of assistive technologies,
- Reliable access to digital platforms,
- Long-term viability of accessibility initiatives.

Robust infrastructure ensures the sustainability of accessibility efforts over time.

10.5. Dimension 5 – Capacity Building

Capacity building encompasses:

- Training for faculty,
- Digital literacy training for students,
- Workshops on accessibility standards and inclusive pedagogy.

Expected outcomes include:

- Improved inclusive teaching practices,
- Enhanced institutional competence,
- Normalization of inclusive digital design.

Capacity building bridges the gap between institutional policy commitments and everyday educational practice.

10.6. Dimension 6 – Policy Integration

Policy integration involves:

- Alignment with Sustainable Development Goal 4 (SDG 4),
- Implementation of WCAG accessibility standards,
- Institution-wide accessibility policies,
- Clear accountability and monitoring mechanisms.

Expected outcomes include:

- Institutionalization of accessibility as a core obligation,
- Measurable and reportable progress in digital accessibility,
- Sustainable, long-term accessibility improvements.

Policy integration anchors digital empowerment within formal governance and quality assurance structures.

11. Future Directions

As digital technologies continue to evolve, the future of accessibility in higher education is expected to be shaped by several emerging directions.

11.1. Context-Aware AI Tutors

Context-aware AI tutoring systems are expected to:

- Respond dynamically to learner behavior and preferences,
- Adapt task difficulty based on performance and cognitive load,
- Provide personalized forms of support when challenges are detected.

Such systems have the potential to support individualized learning pathways for students with visual impairments, enhancing autonomy and engagement.

11.2. Tactile E-Books and Haptic Publishing

Future digital publications are likely to incorporate:

- Tactile diagrams and raised-line graphics,
- Embedded haptic elements,
- audio-based navigation and interaction features.

These innovations may transform how visually impaired students' access, explore, and interact with academic content.

11.3. AI-Driven Data Sonification

AI-driven sonification tools that translate data patterns into auditory cues may enable students to interpret complex datasets through sound. Such approaches hold promise for making scientific and quantitative data fully accessible.

11.4. Blockchain for Accessible Certification

As suggested by Itmazi & Al-Shamesti (2025), blockchain technologies may support:

- Secure and tamper-resistant certification,
- Improved access to academic records,
- Long-term preservation of credentials in accessible formats.

These features may be particularly beneficial for students with disabilities who encounter barriers in accessing or verifying academic documentation.

11.5. National Accessibility Standards

National governments are increasingly expected to:

- Mandate WCAG compliance in higher education,
- Promote accessible digital procurement policies,
- Require transparent reporting on accessibility outcomes.
- Collectively, these measures may normalize accessibility as a standard and expected component of higher education systems.

Conclusion

This chapter conceptualizes digital empowerment as a comprehensive approach to supporting students with visual impairments within higher education contexts. Rather than framing accessibility as a matter of isolated “accommodations” for individual learners, the chapter emphasizes fundamental rights to access, autonomy, and agency in digital learning environments. When universities integrate Universal Design for Learning principles, strategically deploy assistive and emerging technologies, adopt inclusive pedagogical practices, and implement coherent institutional policies, they are better positioned to create and sustain learning environments that respond effectively to the needs of visually impaired students.

The comparative case studies presented illustrate that meaningful and sustainable change emerges primarily at the level of institutional commitment, including policy support, faculty training, and ongoing technological maintenance. The six-dimension framework proposed in this chapter serves as a guide for institutions seeking to approach digital accessibility in a systematic and integrated manner. Ultimately, advancing digital empowerment beyond incremental improvements is not solely a technological challenge but depends fundamentally on institutional choices to actively engage visually impaired learners in the co-creation of inclusive solutions. When accessibility is treated as a core institutional value, higher education becomes not only more inclusive but also more capable of meeting the diverse needs of all learners.

References

Abu Shokhedim, S., Khlaif, Z. N., Hattab, M. K., & Itmazi, J. (2025). Digital empowerment of undergraduate students with vision disability. *British Journal of Visual Impairment*, 0(0). https://doi.org/10.1177/02646196251330161

Al-Azawei, A., & Serenelli, F. (2021). Accessibility and digital inclusion in higher education: A review of global practices. *Australasian Journal of Educational Technology, 37*(3), 38–55.

Ally, M., & Wark, N. (2020). Assistive technologies for inclusive education in the digital era. *Computers & Education Open, 1*(4), 100045.

American Academy of Ophthalmology. (2022). *Night blindness: Symptoms and causes*.

American Foundation for the Blind. (n.d.). *Low vision and legal blindness: Terms and descriptions*. https://afb.org/blindness-and-low-vision/eye-conditions/low-vision-and-legal-blindness-terms-and-descriptions

Birch, J. (2012). Worldwide prevalence of red–green color deficiency. *Journal of Optometry*.

Bourne, R., et al. (2021). Trends in prevalence of blindness and vision impairment. *Lancet Global Health*.

Cleveland Clinic. (2024). *Night blindness (nyctalopia): What it is, causes & treatment*. https://my.clevelandclinic.org/health/symptoms/10118-night-blindness-nyctalopia

Color Blindness Awareness. (n.d.). *Types of color blindness*. https://www.colorblindawareness.org/color-blindness/types-of-color-blindness/

Cook, J. S., & Cook, L. (2004). *The ethics of Web design: Ensuring access for everyone*. Idea Group Inc.

Daniels, S., & Scott, M. T. (2020). Barriers to higher education for blind and low-vision students: Faculty perspectives. *Journal of Higher Education Policy and Management, 42*(2), 134–148.

Envision. (2024). *Envision glasses: AI-powered smartglasses for the blind and visually impaired*. https://www.letsenvision.com/glasses/home

Hartley, J. S. (2018). The role of assistive technology in promoting academic success for college students with visual impairments. *Assistive Technology Outcomes and Benefits, 12*(2), 55–67.

Itmazi, J., & Al-Shamesti, O. (2025). Blockchain and inclusive education in the Arab world: Toward equitable digital transformation. *Journal of Smart Learning Environments, 12*(1), 55–72.

Kelly, S., & Smith, L. (2021). Inclusive education and the digital divide: Accessibility challenges for visually impaired learners. *Disability & Society, 36*(5), 611–629.

Khumalo, N., & Naidoo, S. (2024). Institutional adoption of assistive technologies in African universities. *African Journal of Education and Technology, 18*(1), 1–15.

Kouroupetroglou, G., Shi, Y., & Pfeiffer, T. (2023). Multisensory learning and tactile interaction for visually impaired students. *International Journal of Inclusive Education, 27*(4), 812–829.

Lazar, J., Goldstein, D., & Taylor, A. (2015). *Ensuring digital accessibility*. Morgan Kaufmann.

Lazar, J., Goldstein, D., & Taylor, A. (2015). *Ensuring digital accessibility: A practical guide to universal design.* Morgan Kaufmann.

Lohse, R., Meyer, A., & Gordon, D. (2022). Visual impairment and accessibility barriers in digital learning environments. *Journal of Educational Research, 95*(3), 287–299.

Lueck, A. H., & Dutton, G. N. (2015). *Vision and the brain: Understanding cerebral visual impairment in children*. AFB Press.

MedlinePlus. (2024, August 5). *Blindness and vision loss*. https://medlineplus.gov/ency/article/003040.htm

Meyer, A., Rose, D. H., & Gordon, D. (2014). *Universal Design for Learning: Theory and practice*. CAST Publishing.

Muzenda, A., & Sibanda, P. (2023). Inclusion and accessibility in online education: Evidence from Zimbabwean universities. *Journal of Inclusive Digital Education, 3*(2), 44–59.

National Center for College Students with Disabilities (NCCSD). (2021). *Higher education accessibility reports*.

National Eye Institute. (2023). *Types of color vision deficiency*. https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness/types-color-vision-deficiency

OpticalImages. (2021, December 23). *What's the difference between low vision and being legally blind?* https://www.opticalimages.com/whats-the-difference-between-low-vision-and-being-legally-blind/

Petrie, H., & Bevan, N. (2009). The evaluation of accessibility, usability and user experience. In *The universal access handbook*.

Pfeiffer, T., Shi, Y., & Kouroupetroglou, G. (2022). Emerging technologies for assistive learning: From AI to haptics. *Educational Technology Review, 41*(1), 88–107.

Pfeiffer, T., Shi, Y., & Kouroupetroglou, G. (2022). Emerging technologies for assistive learning: From AI to haptics. *Educational Technology Review, 41*(1), 88–107.

Pune Blind School. (2024). *Top devices for visually impaired individuals*. http://puneblindschool.org

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Selwyn, N. (2020). *Education and technology: Key issues and debates*. Bloomsbury Academic.

Suciu, C.-I., Perju-Dumbrava, L., Nicoara, S. D., et al. (2020). Today's burden affecting the quality of life for color-blind patients. *Romanian Journal of Neurology, 19*(2). https://doi.org/10.37897/RJN.2020.2.1

Tadele, B., & Mengistu, T. (2024). Digital empowerment training for visually impaired students in Ethiopia. *African Education Review, 21*(1), 99–116.

UNESCO. (2023). *Global report on digital inclusion in higher education*.

UNESCO. (2023). *Global report on digital inclusion in higher education.*

UNESCO. (2024). *AI and accessibility: Reimagining inclusive digital education*.

W3C – Web Accessibility Initiative. (2018). *Web Content Accessibility Guidelines (WCAG) 2.1*.

World Health Organization. (2023, August 10). *Blindness and vision impairment*. https://www.who.int/news-room/fact-sheets/detail/blindness-and-vision-impairment

Appendices

Appendix A :Expanded Classification of Assistive and Empowerment Technologies

This appendix presents a structured and academically refined classification of assistive technologies designed to support students with visual impairments in higher education. The classification expands the original framework by introducing clearer categories, explicit educational impacts, and representative use cases.

A.1 Low-Tech Tools

Technology	Description	Educational Impact	Use Cases
Magnifiers	Handheld or stand based optical tools that enlarge printed text.	Enable students to read printed materials, labels, and documents more effectively.	Libraries, exams, printed textbooks.
Tactile Maps & Diagrams	Raised-line graphics that represent shapes, layouts, or spatial locations.	Support conceptual understanding through tactile exploration.	Geography, anatomy, introductory engineering.
Large-Print Materials	Printed resources presented in enlargedfonts.	Reduce visual strain and improve reading accuracy during academic tasks	Examinations, worksheets, reading tasks.

A.2 Mid-Tech Tools

Technology	Description	Educational Impact	Examples
Talking Calculators	Devices that verbally announce numbers and mathematical operations.	Support independent engagement in mathematical tasks.	Algebra, statistics.
Digital Audio Recorders	Tools used to capture and store lecture audio.	Enable learners to review instructional content at their own pace.	Lecture- intensive courses.
Portable Note-Takers	Braille or speech-based devices for digital note-taking.	Support organized note-taking without reliance on visual input.	Seminars, group work.

A.3 High-Tech Solutions

Technology	Functional Role	Examples	Educational Impact
Screen Readers	Convert on-screen text into speech or Braille or Braille output.	JAWS, NVDA	Enable full access to digital interfaces.
OCR Tools	Convert printed or scanned text into machine-readable digital formats.	Seeing AI, Google Lookout	Make inaccessible PDFs usable.
Refreshable Braille Displays	Tactile device that displays dynamic Braille characters.	BrailleNote Touch+	Allow reading and writing in class.
AI Assistants	Provide summaries, explanations, and navigational assistance.	ChatGPT, Alexa	Enhance autonomy and digital confidence.

A.4 Wearables and IoT Devices

Device	Function	Key features	Educational benefits
OrCam MyEye	Miniature camera equipped with AI-based recognition.	Reads printed text and identifies faces and objects.	Enhances classroom participation and access to visual information.
Envision Glasses	AI-powered smart glasses for visual interpretation	Real-time text recognition and object identification	Supports independence mobility and environmental awareness.
Aira Smart Glasses	Remote visual interpretation provided by trained human agents	Live audio guidance through a connected device.	Assist with navigation and completion of complex academic tasks.

A.5 Haptic & Immersive Tools

Technology	Description	Examples	Academic Benefits
Haptic Feedback Wearables	Devices that translate sound into tactile vibrations	NeoSensory Buzz	Improve environmental awareness and sensory substitution.
Immersive Gloves	Wearable devices that reproduce textures and spatial forms.	HaptX Gloves	Enable tactile exploration of STEM concepts and structures.
VR/AR Tactile Environments	Immersive learning environments designed for non-visual interaction.	Touching the Invisible	Facilitate understanding of abstract and spatial concepts.

A.6 Emerging Technologies

Technology	Description	Educational Promise
Brain–Computer Interfaces (BCIs)	Systems that translate neural signals into digital commands.	Potential for autonomous navigation and alternative learning interaction.
AI Multimodal Systems	Technologies that interpret images, text, audio, and gestures	Deliver highly adaptive and personalized learning support.
Blockchain- based Records	Secure, accessible systems for academic certification	Enable reliable verification and long-term accessibility of credentials for visually impaired students.

Appendix B: **Enhanced Conceptual Framework for Digital Empowerment**

This appendix outlines an expanded conceptual model integrating six interrelated dimensions: Access, Autonomy, Agency, Infrastructure, Capacity Building, and Policy Integration.

B.1 Textual Representation of the Framework

Inputs:

- ICT infrastructure,
- Availability of assistive technology,
- Faculty and learners training,
- Accessibility policies.

Processes:

- Adoption of assistive and AI technologies,
- Integration of UDL principles,
- Cross-unit collaboration,
- Accessible digital content design.

Outputs:

- Increased learner autonomy,
- Expanded access to instructional materials,
- strengthened agency and participation,
- Improved academic outcomes.

Feedback Loop:

- Continuous evaluation,
- Data-driven refinement,
- Student-centered redesign,
- Policy refinement.

B.2 Explanatory Narrative

1. Access ensures equitable engagement with digital platforms.
2. Autonomy emphasizes independent navigation supported by AI and multisensory tools.
3. Agency foregrounds learner participation in co-creating solutions.
4. Infrastructure sustains accessibility through maintenance and robust networks.
5. Capacity building develops inclusive competencies among faculty and students.
6. Policy integration institutionalizes accessibility as a mandatory, accountable practice.

Appendix C: Comprehensive Summary of Case Studies

This appendix synthesizes evidence from Palestine, Ethiopia, South Africa, Canada, and Finland.

C.1 Comparative Table

Region	Key Study	Findings	Implications
Palestine	Abu Shokhedim et Al. (2021)	AT improved GPA, autonomy, participation	Need for Arabic AT tools and trained staff
Ethiopia	Tadele & Mengistu (2024)	DET increased digital literacy, confidence	Structured training essential
South Africa	Khumalo & Naidoo (2024)	Governance + funding determine AT adoption	Sustainability matters
Canada	UNESCO (2023)	AI libraries ↑ access by 40%	Government-driven inclusion works
Finland	UNESCO (2024)	UDL in national policy	Systemic design ensures consistency

C.2 Analysis

Across diverse regions, a pattern emerges:

- Training → Higher confidence,
- Policy → Consistent accessibility,
- Governance → Sustainability,
- Localization → Effectiveness,
- Funding → Long-term adoption.

These elements form global pillars of digital empowerment.

Appendix D: Practical Guidelines for Universities

This new appendix offers clear, actionable guidelines for institutions aiming to enhance accessibility for visually impaired learners.

D.1 Guidelines for Faculty

Provide alternative text; audio-described media; accessible PDF/DOCX materials; avoid color-only cues; share notes in advance; apply UDL in course design.

D.2 IT Departments

Ensure WCAG 2.1 compliance; conduct regular accessibility audits; enable screen-reader installation; supply accessible templates.

D.3 University Leaders

Establish accessibility offices; allocate sustainable budgets; embed accessibility in accreditation; monitor progress via measurable indicators.

Appendix E: Recommended Tools and Applications

This appendix presents a curated list of tools that visually impaired students can use in everyday academic tasks.

E.1 Reading and Text Access

- NVDA,
- JAWS,
- Seeing AI,
- Kurzweil 1000,
- Voice Dream Reader.

E.2 Productivity and Organization

- Google Assistant,
- ChatGPT Voice Mode,
- Notetaker,
- BrailleNote Touch+.

E.3 Navigation and Daily Living

- Aira,
- Be My Eyes,
- Envision AI,
- BlindSquare,
- 3D-printed tactile models,
- HaptX Gloves,
- TactileView.

Chapter 3

Strategic Leadership for School Improvement: The Malaysian Experience

Muhammad Faizal A. Ghani*
Norfariza Mohd Radzi
Nur Diyana Zakariah
and Amirul Fahmie Abdul Razak
Faculty of Education, Universiti Malaya, Malaysia

Abstract

School improvement in Malaysia continues to be challenged by disparities in student achievement, resource limitations and the pressures of systemic reform. Those circumstances highlight the necessity of a leadership that goes beyond the role of procedures to include strategic vision, flexibility and transformational ability. The chapter explores how strategic leadership has implemented school improvement in Malaysia education system. It initially explains the meaning of strategic leadership by contextualizing it in the context of international scholarship and national policy agenda. Reform agendas like the Malaysia Education Blueprint 2013-2025 have put increased pressure on school leaders to be foresighted and resilient as well as results-oriented practitioners. In order to address such demands, the chapter introduces the Strategic Leadership Model for Malaysian School Leaders (STALEM), which comprises eight domains: (i) leading change strategies, (ii) cultivating futurist thinking, (iii) advancing shared vision and values, (iv) fostering personal growth, (v) strengthening stakeholder collaboration, (vi) sustaining conducive organisational climates, (vii) ensuring equitable power distribution and

* Corresponding Author's Email: mdfaizal@um.edu.my

In: Progress in Education. Volume 91
Editor: Roberta V. Nata
ISBN: 979-8-90134-375-3

(viii) driving competency development programmes. These domains provide pathways for mobilising collective capacity and addressing contemporary educational challenges. Drawing on Malaysian policy priorities and empirical research, the chapter demonstrates how strategic leadership enhances teaching quality, strengthens stakeholder partnerships and cultivates professional learning cultures. It also determines the obstacles encountered by leaders such as lack of resources and conflicts between centrally-devised policy systems and local environments. The chapter ends with implications to policy and practice including leadership development, professional training and governance reform as key to further improvements in the provision of education of quality and equity.

Keywords: strategic leadership, school improvement, Malaysian education, educational leadership, STALEM model

1. Introduction

The problem of school improvement has been one of the primary issues in the research of education. Hopkins (2001) claims to achieve sustainable improvement, systemic change should be used instead of piecemeal interventions and Fullan (2014) indicates that reform can only be successful when leadership mobilises people, processes and resources around an explicit moral purpose. Such views place leadership at the centre of entrenching reform within the school structures in way that change becomes sustainable and significant.

Malaysian reforms over the years have influenced the nature of the education quality and equity concerns by placing student achievement, teacher professionalism and systemic transformation as the ambitious goals of the education system (Ministry of Education Malaysia, 2013). Ko et al. (2012) emphasize that these reforms in Asia make things get better but also demonstrate contradictions between the directives of the top-down policy and the reality on the school level. The same is noted by Cheng (2019), who argues that schools in the area, especially in the multi-ethnic and diverse environment of Malaysia, tend to have issues balancing between national concerns and the needs of the local population. Yet despite these ambitious reforms, deep-rooted challenges persist.

Nevertheless, problems continue to persist even with the reform efforts. According to Selvaratnam (2022), the results of students in the country have been continuing to be uneven with differences in urban and rural schools, as well as socio-economic and ethnic boundaries. Siti Noor Ismail et al. (2020) further contends that the ability to manage multicultural school communities, shortage of teachers and resource constraints only adds to the strains on school leaders. These difficulties corroborate the fact that technical reforms are not enough. Leithwood et al. (2020) underline that leadership comes second after classroom teaching in the effects it has on student learning and thus it is the most crucial factor in improving the capacity of schools.

This recognition has shifted attention from administrative leadership toward strategic leadership. According to Bush (2021), leadership should not be administrative as it should be strategic, which allows the leader to predict the future needs, match the resources with the long-term objectives and develop professional cultures that drive the transformation. The same idea is supported by Harris et al. (2017) who provide the information that Malaysian principals are anticipated to become more and more visionary and innovative agents of change instead of merely adhering to bureaucratic demands. Jäppinen (2017) also confirm the trend in the educational discourse worldwide, where leadership is viewed not as a managerial task but as a sense-making process, as a direction-setting and capacity-building process.

Building on these perspectives, this chapter explores how strategic leadership can be used to foster the school improvement in the Malaysian education system. The chapter is organised into seven sections. The discussion begins by reviewing theoretical perspectives and models of strategic leadership, before situating the concept in the Malaysian context by examining its evolution, expectations and challenges. The chapter then presents the derivation of the STALEM framework consisting of eight areas that are interrelated and explores its practical applications. It concludes with implications for policy and leadership development and reflects on the contribution of STALEM to school improvement in Malaysia.

This chapter uses a connection between theory, policy and practice to argue that strategic leadership has a role to play in making the Malaysian dream of education a reality. The analysis proves that strategic leadership can not only improve the quality of teaching, reinforce the involvement of stakeholders and promote cultures of professional learning but also deal with the structural and cultural obstacles that leaders have to overcome.

Despite the prevalence of research on strategic leadership at the international level, there is limited literature that explores the application of strategic leadership in Malaysia in its context of combining global theoretical viewpoint with the national policy-making systems. The chapter thus adds the novelty of introducing the Strategic Leadership Model of Malaysian School Leaders (STALEM) as a model that is both practical (local) and internationally familiar. Through this, it provides new knowledge to scholars, policy makers and practitioners on how strategic leadership may be mobilised to promote school improvement across a variety of systems that are reform oriented.

2. Theoretical Perspectives on Strategic Leadership

The education field is dynamic, and it is constantly undergoing change with regards to social change, global pressures and national desires. The role of leadership is inevitable in such an environment to lead schools through reform and transformation. The achievement of any education system is not only based on policy framework but also the ability of the leaders to foresee challenges and put on the shoulder of the internal resource to match with external needs. Strategic leaders need to be in a position to decipher the external environment, as Stewart and Shamdasani (2014) state, but they should also marshal the internal resources of their organisations.

Educational leaders are thus expected to make strategic thoughts so as to steer their schools to excellence. As stressed by Goldman et al. (2015), strategic level leaders are expected to build wisdom, the capability to think creatively and have the capacity to connect and synthesise various ideas. These attributes make strategic leadership a unique process to ordinary administration as they are visionary and dynamic and define the orientation of organisations.

The current chapter analyses the notion of strategic leadership in the educational field by focusing on a number of critical areas. It starts with an examination of the concept of leadership in general and then proceeds to examining the concept of the leader and what strategy entails. It offers a basis on which strategic leadership can be defined and its application in the education field can be thought about. This is followed by the discussion of theories of strategic leadership which provide conceptual basis of practice and research. These factors collectively create the theoretical foundation on which the chapter constructs its research on strategic leadership in the Malaysian environment.

2.1. The Concept of Leadership

The concept of leadership continues to evolve due to its context, objective and domain, which may include relationships, competencies and skills. Cowley (1928) offered one of the earliest formal definitions of leadership, stating that it entailed an office role within an organisation. Since Cowley, more leadership definitions have emerged, advancing from the more systematic approaches with contextual relevance (Stogdill, 1974).

According to Yukl (2012), leadership is one of those notions that is subjective and variable. In contrast, Bass (2007) characterises leadership as a process of significance, adding that it is an ever-changing phenomenon which is not intrinsic, existing only in the relationship of leaders and followers. In the same spirit, Northouse (2021) further insists that leadership must be seen and exercised as an interactive process from both parties, suggesting that it is not mere one-directional influence as the leader to the followers. This view highlights the interdependence of the followers and the leaders as well as the relational quality of leadership.

Bush (2020) identifies three dimensions in the study of leadership which are, leadership as influence, leadership and values and leadership and vision. These dimensions show clearly that leadership is and ought to be transformative, for to lead is to change the direction and culture of an organisation. Other authors have associated the initiation and maintenance of change with leadership (Davies, 2009; Fullan, 2004), the influence of others to the attainment of common objectives (Stogdill, 1974; Ibrahim Mamat, 1998; Yukl, 2012) and the unification of values that control organisational activities (Wasserberg, 1999). In the context of reform, Joyce (2012) in Strategic Leadership in the Public Service, argues that leadership in the public sector becomes more important when governments are seeking transformative change.

From these definitions, it can be concluded that leadership encompasses the skills and capacities of individuals to influence subordinates toward achieving organisational objectives. Leadership comprises two central elements: the goals to be achieved and the means of attaining them. Goals reflect the vision and mission established by the organisation, while the operative means often involve persuasion, influence and the mobilisation of collective effort. In educational settings, leadership can therefore be understood as the capacity of school leaders to guide their teachers, staff and communities through processes of planned change aimed at fulfilling the future-oriented mission and vision of schools.

2.2. The Concept of Leader

In the field of education, leaders bear considerable responsibility in sustaining the goals and agenda of schooling, particularly within complex global settings. In this regard, Drucker (1996) posits that the presence of a leader can only be explained by the presence of followers, whereas Gardner (2011) includes that leaders in the educational context must also serve as personal exemplars and shape the attitudes and behaviours of their constituencies.

Fiedler (1967) characterised leaders as persons whose responsibilities include overseeing a group of people in the pursuit of objectives. In terms of strategy, Hallinger (2011) emphasises the need for leaders to appreciate their institutional context. Additional clarity is provided by Kotter's (1996) model of the leadership process, in which leaders articulate a vision and strategy, communicate that strategy, inspire and energize people and realize significant transformations. This view is reinforced by Gill (2013), who describes leadership as an endless cycle of directing, aligning and maintaining focus.

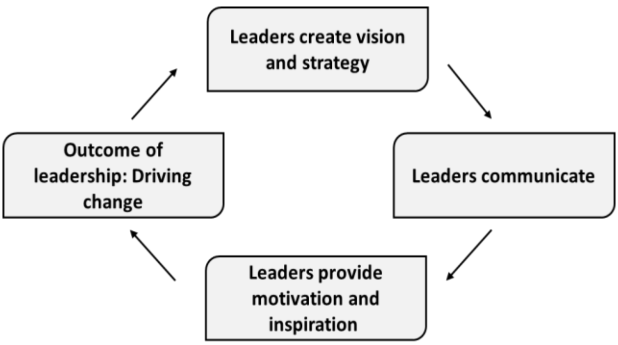

Figure 1. Leadership process (adapted from Gill, 2013).

Educational institutions confer formal leadership roles to individuals such as directors, principals and head teachers. Beyond administrative functions, leadership entails crafting vision, spearheading change, inspiring individuals and fostering a positive ecosystem. Leaders assist in the support and organisation of the group, procedural formation, facilitating shared decision-making and creating environments conducive to professional development. In the context of schools, this involves the integration of the planning, organisation, supervision and evaluation of all educational activities in such a way that the purposes of teaching and learning and the goals of the community, are harmonised.

2.3. The Concept of Strategy

The concept of strategy, though defined in various ways, remains fundamentally tied to planning. The term derives from the Greek *strategos*, originally referring to a general commanding an army (Gill, 2013). Military strategy was characterised by command-and-control and a top-down approach. In Malay usage, the *Kamus Dewan* defines strategy as a systematic plan that accounts for multiple factors in order to achieve goals. In management literature, Chandler (1962) was influential in defining strategy as the determination of long-term goals and objectives, along with the actions and resource allocations needed to achieve them.

By the 1990s, the understanding of strategy had expanded beyond long-term planning to include alignment with dynamic external environments and the integration of leadership considerations (Gill, 2013). Hambrick and Fredrickson (2005) further emphasise that strategy is shaped through strategic analysis and clearly articulated objectives that reflect the organisation's underlying values and purpose (see Figure 2).

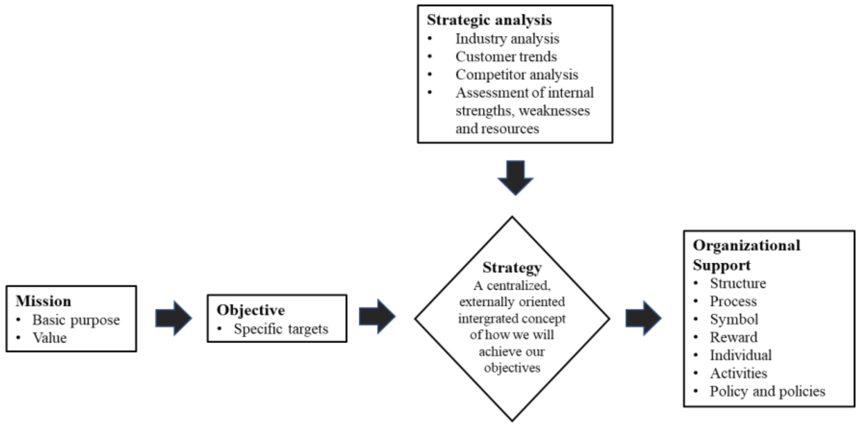

Figure 2. Strategy in leadership (adapted from Hambrick & Fredrickson, 2001).

In education, strategy refers to deliberate planning aimed at addressing stakeholder needs while remaining consistent with institutional missions and visions. Although the concept appeared in educational management literature in the 1980s, more systematic research emerged in the 1990s (Eacott, 2008a, 2008b). Today, strategy in education is understood as holistic and intentional planning that guides schools in meeting evolving challenges.

2.4. The Concept of Strategic Leadership

Strategic leadership has emerged as a crucial capability for organisations seeking development and competitive advantage. Ireland and Hitt (1999) define it as the ability to anticipate, envision, remain flexible, think strategically and collaborate with others to initiate change that secures the future. Boal and Hooijberg (2000) emphasise three elements of strategic leadership: the capacity to absorb change, to effect change and to apply managerial wisdom.

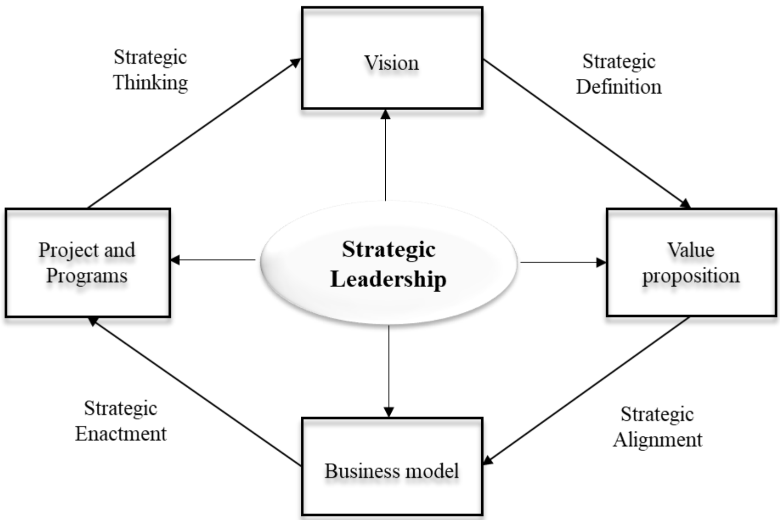

Figure 3. Strategic leadership and the cycle of strategy (adapted from Finkelstein et al., 2009).

Strategic leadership is oriented toward direction-setting through purpose, vision and communication (Alan & John, 2001; Adair, 2005). Hambrick (1989) adds that it involves aligning organisations with external forces such as technology, markets and regulations. From a governance perspective, Finkelstein et al. (2009) argue that strategic leadership is reflected in how top leaders shape organisational vision and outcomes through cycles of strategic thinking, definition, alignment and enactment.

This view is reinforced in their model of strategic leadership as a cycle of four interrelated processes: strategic thinking, strategic definition, strategic alignment and strategic enactment (Figure 3). Strategic thinking entails holistic analysis to generate vision, strategic definition involves refining the organisational value proposition, strategic alignment requires configuring the business model to deliver this value and strategic enactment consists of implementing the necessary projects and programmes.

Recent perspectives emphasise adaptability and innovation. Luciano et al. (2020) describe strategic leadership systems as involving boards and top management teams that work interdependently. Suharto (2023) highlight strategic leadership as an ongoing process of analysis, planning, implementation, monitoring and evaluation. Funda (2022) presents it as a fusion of innovation and transformational leadership, while Al Huraizi and Marni (2023) stress anticipation and flexibility. Jabbar and Hussein (2017) argue that strategic leadership functions at the top levels of organisations to shape strategic direction.

Taken together, strategic leadership is not confined to individuals at the apex of structures. Rather, it is enacted by those who can interpret external challenges, drive innovation and sustain resilient organisational cultures.

2.5. The Concept of Strategic Leadership in Education

Strategic leadership is important to the organisational climate and its persistence in ensuring a continuous growth of quality in the education institutions. According to Suharto (2023), strategic leadership has been promoted as a useful leadership strategy and it will always feature in effective implementation of organisational strategies. Equally, it is one of the five most popular methods of leadership that Neyişci and Erdoğan (2022) state have become credible in theory and practice in the current educational management. Boal and Hooijberg (2000) stress on leaders capable of creating and developing organisational capacity as the strategic leadership in education and Davies and Davies (2004) put it as the development and implementation of change and improvement strategies that lead to effectiveness and excellence in schools. Pisapia (2009) goes on to add to this view by introducing strategic leadership as a practical and practice based substitute in the interpretation and implementation of leadership in education. Eacott (2011) helps us remember that the contemporary educational leaders are forced to manoeuvre through the presence of historical, cultural, social and political complexities, unstable economic situations, blistering development of technologies and more diverse groups of students.

In a world that is changing at a rapid pace, especially due to the digital revolution, educational leaders should be progressive and attentive. Nur Hidayah et al. (2015) emphasise that leaders should be informed of the technological advancement and the consequences of technological advancement on the organisational operations. Fazleen Mohamad and Siti Noor Ismail (2018) present a similar definition of strategic leadership in education, which is the ability to create and communicate vision and mission and successfully mobilize the available human resources to accomplish the long-term educational objectives. In this framing, strategic leadership in education can be realized as the capacity of the leaders to set direction by embracing the future orientation of futurist and following the vision with the new education emergent demands. It includes spearheading strategic change, initiating competency development programmes to enhance teacher capacity, mobilizing student potential and it creates an organisational climate that is supportive and favourable. In addition, strategic educational leaders are supposed to practice self-leadership, share power equally, duties and responsibly, create consensus with stakeholders and foster effective collaborative cultures in schools. These practices make sure that strategic leadership not just can be viewed as an administrative position, but a transformative process which places schools in long term excellence.

2.6. The Theories of Strategic Leadership

Theories of strategic leadership provide conceptual foundations that guide research and practice. Four prominent models are reviewed here: Ireland and Hitt's (1999) Strategic Leadership Theory, the 5P Model (Pryor et al., 2007), the Strategic Leadership Model (Davies & Davies, 2004) and the Integrated Strategic Leadership Model (Boal & Hooijberg, 2001). Each theory highlights different dimensions of strategic leadership, ranging from anticipatory capacity and adaptability to structural alignment and personal characteristics. Together, they illustrate how leadership functions as both an organisational and individual capability, shaping the ability of schools to respond to reform, innovation and accountability.

2.6.1. Strategic Leadership Theory (Ireland & Hitt, 1999)

Ireland and Hitt's (1999) theory of strategic leadership is among the most influential frameworks for understanding how leaders can ensure organisational success in the turbulent contexts of the twenty-first century. They define strategic leadership as the ability to forecast, envision, remain flexible, think strategically and work with others to initiate change that will create a viable future for the organization. This understanding highlights leadership as not just a position of power but rather a dynamic capability to involve people and resources towards long term strategic objectives.

According to this theory, effective strategic leadership is characterised by three interrelated roles: setting a clear strategic direction, establishing balanced organisational controls and sustaining an effective organisational culture. Through these roles, leaders are able to align organisational behaviour with external demands, while simultaneously exploiting new opportunities for growth and improvement. Ireland and Hitt (1999) argue that what differentiates organisations in the twenty-first century is not only their structures or resources but also how top leaders enact their strategic leadership responsibilities in interaction with senior management teams and organisational members.

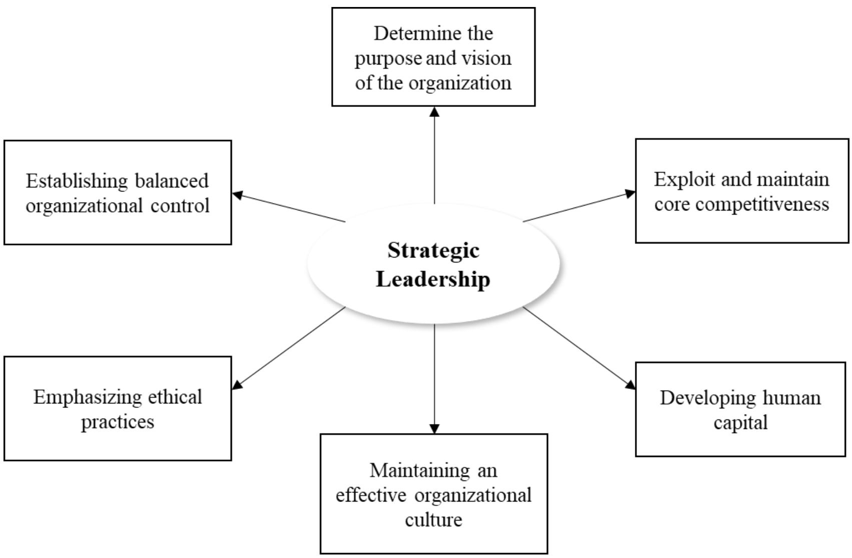

Figure 4. The Strategic Leadership Theory by Ireland & Hitt (1999): Six Practices of Effective Strategic Leadership.

The model identifies six core practices of effective strategic leadership (see Figure 4), which revolve around the integration of diverse competitive information, the formulation of adaptable strategies and the use of broad-based strategic controls. While originally developed in the corporate domain, the theory has significant relevance for education. Schools, much like corporations, operate in highly complex environments where leaders must anticipate change, balance accountability systems and sustain cultures of learning that promote long-term improvement.

In the educational context, the application of Ireland and Hitt's framework highlights the necessity for school leaders to integrate multiple streams of information, balance national policy demands with local needs and cultivate a culture that supports both innovation and stability. Strategic leadership, therefore, becomes the bridge between external reform pressures and internal organisational capacity, positioning leaders as agents of sustainable transformation in schooling.

This theory highlights the capacity of leaders to anticipate, envision and maintain flexibility while empowering others to create strategic change. In the context of schools, this translates into principals anticipating future educational demands, envisioning long-term improvement and empowering teachers and staff to innovate. It reminds us that school leadership is not simply reactive, it requires foresight and adaptability to sustain meaningful reform.

2.6.2. The 5P Strategic Leadership Model (Pryor et al., 2007)

The 5P Strategic Leadership Model by Pryor, Anderson, Toombs and Humphreys (2007) is a very useful model to comprehend how strategic and tactical aspects of leadership interact. This model underlines the fact that successful leadership demands the decisions and the actions that are at once strategic and operational and involve a combination of various aspects of organisational operation in a logical way. The model distinct five factors that are interdependent and are the foundation of strategic leadership:

1. Purpose – the strategic intent of the organisation, encompassing mission, vision, goals and strategies;
2. Principles – the operational guidelines and core values that underpin organisational culture and behaviour;
3. Processes – the systems, structures, procedures, infrastructure and rules that support organisational functioning;

4. People – the individuals and teams who are engaged in processes and responsible for implementation; and
5. Performance – the measurable outcomes of organisational work, typically captured through key performance indicators and results.

As illustrated in Figure 5, the arrows demonstrate the causal relationship among these five elements. Strategy (Purpose) drives the internal and external structures (Principles and Processes), which in turn shape the behaviour of individuals and teams (People), ultimately leading to outcomes (Performance). Importantly, the feedback loop from Performance back to Purpose ensures continuous alignment between organisational outcomes and strategic objectives. This cyclical mechanism highlights the centrality of monitoring and feedback in sustaining effective strategic leadership (Pryor et al., 2007; Kabeyi, 2019).

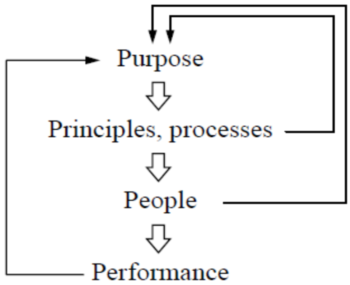

Figure 5. The 5P Strategic Leadership Model (Pryor et al., 2007).

Although originally designed for the corporate management context, Pryor et al. (2007) emphasise that the principles of this model are applicable across domains, given the overlapping nature of leadership and management. Within the field of education, the 5P model offers a valuable lens for school leaders who are required to engage in strategic management processes. For example, educational leaders must articulate a clear vision and mission (Purpose), establish a strong value-driven culture (Principles), design effective organisational and pedagogical systems (Processes), empower teachers and staff (People) and evaluate outcomes based on school performance and student achievement (Performance).

The 5P model a comprehensive framework for aligning organisational strategy. In educational settings, this can be operationalised by clarifying the school's purpose, embedding shared values, ensuring effective processes, investing in teacher capacity, and monitoring outcomes. It underscores how leadership in schools must integrate moral purpose with practical management to secure improvement. By adopting the 5P model, school leaders can ensure that strategic quality management is embedded in educational practice, linking purpose to performance in a continuous cycle of improvement.

2.6.3. The Strategic Leadership Model (Davies & Davies, 2004)

The Strategic Leadership Model proposed by Barbara J. Davies and Brent Davies (2004) was developed in the context of education, making it highly relevant and applicable to school leadership. This model emphasises two fundamental dimensions of strategic leadership: the leader's capacity for organisational management and the leader's personal characteristics. Together, these dimensions form the basis for effective leadership in educational settings.

Organisational Management Capacity

Davies and Davies (2004) argue that strategic leaders must demonstrate a range of managerial capacities in order to lead schools effectively. These include:

- Strategic Orientation: the ability to adopt a long-term perspective, anticipate the future and develop a strategic vision for the school (Stacey, 1992; Adair, 2002; Beare, 2013). Strategic orientation requires leaders not only to plan for current needs but also to establish a roadmap for desired future outcomes.
- Translation of Strategy into Action: the ability to operationalise strategic ideas into concrete actions. Leaders must restructure the organisation, identify critical projects and mobilise resources in ways that align the school with its future direction (Tichy & Sherman, 1993).
- Organisational and People Alignment: the capacity to align staff and organisational structures with the strategic vision. This involves fostering commitment through shared values and building a strong organisational culture (Boal & Bryson, 1988; Gioia & Thomas, 1996).
- Identification of Effective Strategic Interventions – the skill to determine the right points of intervention for implementing change.

Leaders must not only create new strategies but also identify the right time and method for executing significant changes.

- Development of Strategic Competence: the ability to build core capabilities that enable the school to reconfigure existing resources and adapt to new challenges (Prahalad & Hamel, 2009).

Personal Characteristics of Strategic Leaders

Beyond managerial skills, Davies and Davies (2004) highlight the personal dispositions that define effective strategic leaders. These include:

- Restlessness: a sense of dissatisfaction with the status quo, reflecting what Senge (1990) terms "creative tension." Strategic leaders are motivated to seek continuous improvement and are willing to challenge existing practices.
- Absorptive Capacity: the ability to assimilate new knowledge, reorganise information and apply it productively within the organisation (Boal & Hooijberg, 2001).
- Adaptive Capacity: the ability to adjust flexibly to complexity, uncertainty and change. Leaders must be responsive and proactive in navigating emerging opportunities and challenges.
- Wisdom: the capacity to exercise sound judgment by balancing competing priorities, responding to environmental pressures and applying knowledge for the collective good.

This model has been widely used in school-level research and practice due to its holistic integration of organisational and personal dimensions of leadership. It provides a useful framework for school leaders who must simultaneously manage institutional systems and embody personal qualities that inspire confidence, adaptability and strategic foresight. This model directly applies to education by positioning leaders as visionaries who cultivate an environment where innovation and collaboration thrive. This suggests that strategic leadership in schools is not confined to planning documents but embedded in the daily culture and practices of the institution. By adopting the Strategic Leadership Model, educational leaders can better align organisational structures with strategic vision, cultivate a culture of shared values and ensure that schools are equipped to respond effectively to the demands of educational reform and innovation.

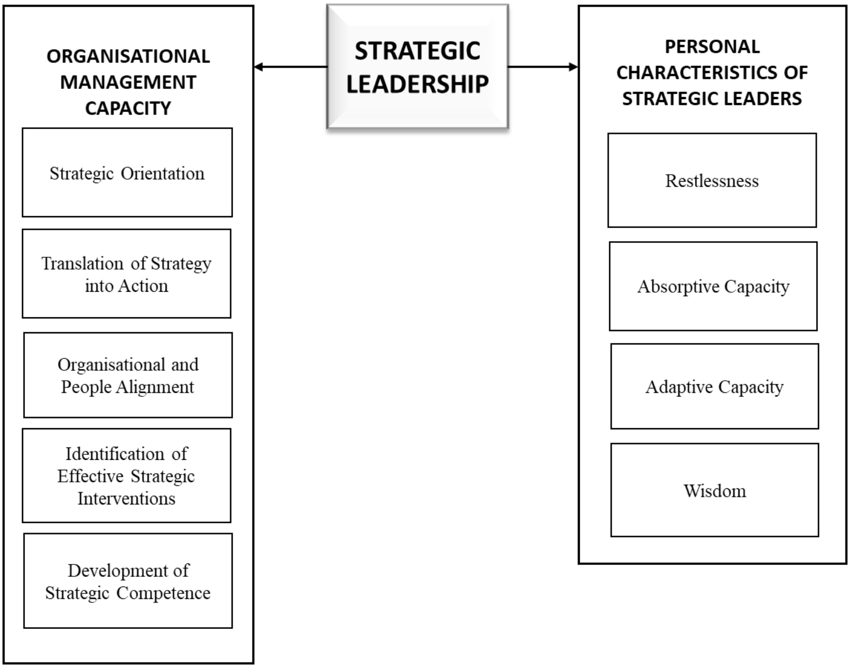

Figure 6. The Strategic Leadership Model (Davies & Davies, 2004).

2.6.4. The Integrated Strategic Leadership Model (Boal & Hooijberg, 2001)

The Integrated Strategic Leadership Model (Boal & Hooijberg, 2001) combines traditional leader- ship with more recent comprehensive cognitive models of visionary, charismatic and transformational leadership. This model illustrates how leadership effectiveness is influenced by the interplay of cognitive complexity, social intelligence and behavioural complexity in conjunction with a leader's ability to adapt, absorb and exercise good judgement. Through the combination of these factors, the model emphasises the importance of strategic leadership in increasing both leadership effectiveness and overall organisational performance.

Strategic leaders, according to Boal and Hooijberg (2001), need to have the capability of gaining, understanding and utilizing new knowledge in environments that are continuously changing. The more absorptive capacity leaders have the more relevant they can remain in complex environments where new information constantly appears. Adaptive capacity is also important because it allows leaders to adjust strategies, processes and styles of leadership in face of external pressures associated with technology changes, market contingencies and social developments. Added to this is the concept of managerial prudence, which describes the ability of managers to use good judgment, make ethical-based decisions and conduct themselves in a way that reflects longer-term goals.

The model illustrates how absorptive capacity, adaptive capability and managerial wisdom interact dynamically to reinforce the effectiveness of strategic leadership. Accordingly, successful leadership is capable of predicting and facing up to new challenges whilst leading their organisations to become sustainable competitors. This is not merely the process of making decisions, but also includes the capacity to excite, motivate and stir followers through vision and transformational influence.

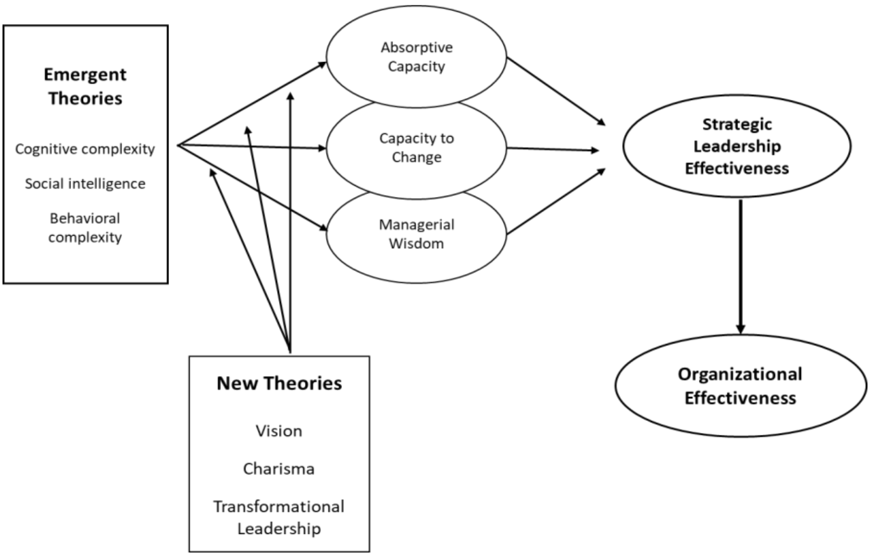

Figure 7. Integrated Strategic Leadership Model (Boal & Hooijberg, 2001).

Boal and Hooijberg (2001) add to this that strategic leaders need to be aware of environmental changes and show courage, but also competitiveness. These leaders need to have a deep knowledge of organizations in order to articulate clear visions and missions that reflect the future. When these personal attributes, charisma and vision articulate with managerial competence you get a full picture of leadership. In practice, this balance also provides that leaders can influence organisational culture, respond to external demands and be effective over time.

As illustrated in Figure 7, the Integrated Strategic Leadership Model emphasises that the combination of absorptive and adaptive capacities with managerial wisdom results in more effective leadership outcomes. These outcomes, in turn, contribute directly to organisational effectiveness by enhancing resilience, promoting innovation and ensuring that the institution remains aligned with its strategic objectives. This model combines managerial systems with the personal attributes of leaders, stressing cognitive complexity, social intelligence and absorptive capacity. In educational contexts, these attributes are essential for navigating policy demands, engaging with diverse stakeholders, and absorbing new knowledge for school improvement. It reflects the reality that effective principals must balance technical management with interpersonal and intellectual skills.

Taken together, these theories and models highlight that strategic leadership is an integrated practice that combines vision, adaptability, collaboration and systemic alignment. Ireland and Hitt (1999) emphasise anticipatory capacity and long-term direction, while Pryor et al. (2007) situate strategy in the interconnected domains of purpose, principles, processes, people and performance. Davies and Davies (2004) foreground the importance of future-oriented thinking and strategic culture in schools, whereas Boal and Hooijberg (2001) stress the interplay between managerial systems and individual leader capabilities. When applied to educational settings, these models collectively suggest that effective school leadership requires both structural planning and human-centred practices, ensuring that reforms are technically implemented and also culturally embedded. This integrative perspective provides a foundation for contextual frameworks which adapts these theoretical insights to the specific challenges and aspirations of Malaysian schools.

2.7. Conceptual Foundations: Strategic Management and Educational Leadership

The theoretical base of strategic leadership is that of intersection of management theory with educational leadership. The strategic management as defined by Chandler (1962) focuses on the long term setting of organisational objectives, the incorporation of courses of action and allocation of resources towards attaining the organisational objectives. Later theorists such as Mintzberg (1994) and Hambrick and Fredrickson (2005) developed this idea by emphasising the sensitivity of organisational structures to environmental needs and that strategy is not a fixed position but is an adaptive response to environmental forces. In this context, strategic leadership was developed as a combination of planning, visioning and adapting.

Strategic leadership in the education sector takes into consideration both management-oriented and leadership-oriented viewpoints. On the one hand, it is based on management principles like efficiency, accountability and resource allocation (Bush, 2020). On the other, it incorporates leadership aspects of impact, collective values and ethical intent (Fullan, 2004). This dual foundation underscores that educational leadership cannot be reduced to technical management alone; rather, it requires an approach that aligns operational effectiveness with aspirational vision and long-term educational purpose.

According to Pisapia (2009), strategic leadership can best be seen as a form of adaptive capacity which allows leaders to move through more intricate systems striking a balance between continuity and innovation. This flexibility is specifically useful in the area of education where leaders have to address the quickly- evolving policy models, the shifting expectations of communities and the varying learning requirements of learners. Eacott (2011) also states that one must not view educational leadership in terms of managerial logics only but rather in terms of the socio-political and cultural contexts according to which the school life is formed. This critical prism places strategic leadership as a technical and ethical practice, which is rooted in larger social discourses.

Accordingly, the sets of concepts of strategic leadership have integrated three strands of interrelated concepts, namely, the managerial orientation to planning and accountability, the leadership orientation to vision and influence and the contextual orientation of acknowledging political, cultural and social realities of education. The three-part framework offers the framework that is required in order to comprehend the conceptualisation and implementation of strategic leadership in schools.

2.8. Defining Features of Strategic Leadership in Schools

Strategic leadership in the schools is characterized by a combination of the theoretical background and a series of practical characteristics that made it outstanding among other leadership styles. According to Williams and Johnson (2013), the characteristics of the strategic leaders in the education sector are visioning, strategy formulation and long-term orientation. These characteristics highlight that leaders need to clearly state the sense of direction of the school, make sure that the mission and practice are always aligned and persist in the anticipation of the future challenges and opportunities.

The second characteristic attribute is that it is able to deal with change proactively. Mincu (2022) states that educational leaders should be able to spearhead change as opposed to responding to external forces. This will come with a need to inspire employees, mobilise assets and maintain a sense of morale during transitional phases. There is also such change leadership, which involves distributed leadership practices, in which power is distributed among teachers, senior leaders and other stakeholders to establish a collective ownership of school improvement (Amels et al., 2020).

Another characteristic is the development of organisational learning. Senge (1990) as well as Perez and Banayo (2023) proposes the idea of the learning organisation which is quite relevant to schools. Strategic leaders stimulate professional growth, support reflection and establish the systems in which feedback can be used to make decisions. Such an orientation supports not only the development of teachers but also introduces the culture of continuous improvement into the school.

Moreover, moral purpose and ethicality is also a unique aspect of education strategic leadership. Day et al. (2016) state that effective school leaders can maintain the process of improvement due to basing their decisions on such values like equity, inclusion and social justice. This moral aspect makes educational strategic leadership unlike the corporate one because it is not only about the success of the organisation, but the comprehensive growth of learners and the society.

Lastly, the ability to be resilient and adaptable is one of the major characteristics of school leadership today. According to the observations mentioned by Al Huraizi and Marni (2023), educational leaders in the wake of technological advances, shifts in policies and the global uncertainties have to be able to anticipate, respond and adapt to rapid change. Strategic leaders should thus develop cognitive flexibility and emotional intelligence to be able to lead with a foresight and empathy.

Collectively, these features support the fact that strategic leadership in schools is multi-faceted. It incorporates vision and direction-setting, proactive change management, distributed leadership, organisational learning, ethical purpose and resilience. These are the defining features that make it an important leadership model in leading schools to sustainable development in more and more complex educational environments.

3. Strategic Leadership in the Malaysian Education Context

Strategic leadership in Malaysia is best understood by situating it within the country's broader educational reforms, cultural diversity and systemic structures. While the concept has gained global prominence as a driver of school improvement, its application in Malaysia reflects specific national imperatives and contextual challenges. Leadership has been positioned as a critical lever in bridging the gap between ambitious reform agendas and the realities of school-level implementation. Over the years, the discourse on leadership has shifted from moral and philosophical foundations to performance-oriented, strategic expectations aligned with international benchmarks. At the same time, Malaysian principals face growing demands to balance accountability, instructional improvement and community engagement within a highly centralised system. The following subsections trace the evolution of leadership discourse in Malaysian education, outline the current expectations of school leaders and examine the challenges they encounter in enacting strategic leadership. Together, these perspectives provide a foundation for understanding why a contextually grounded framework such as the Strategic Leadership Model for Malaysian School Leaders (STALEM) is necessary to strengthen sustainable school improvement.

3.1. The Evolution of Leadership Discourse in Malaysian Education

Educational leadership development in Malaysia has been influenced by the changing national imperatives and global reform agendas. The National Education Philosophy (NEP) which was first implemented in 1988 and revised subsequently, framed the moral and integrated orientation of Malaysian schooling regime with its emphasis on balanced growth in terms of intellectual, spiritual, emotional and physical aspects (Ministry of Education Malaysia, 1988). Muhammad Faizal A. Ghani (2013) suggests that this philosophical bedrock, led us to the concept of school leadership not just as a mechanical role but also a moral enterprise through which the leaders should foster values, harmony and holistic student development.

The Malaysia Education Blueprint 2013–2025 (MEB) marked a turning point in the discourse on leadership by explicitly linking school improvement to leadership capacity. The Blueprint highlighted leadership as a lever for reform and encouraged visionary, instructional and results-oriented principals (Ministry of Education Malaysia, 2013). According to Abdul Rahim Abdul Rahman et al. (2020), this represented a movement away from former managerial paradigms of leadership to a more strategic model of leadership, which reflected national performance targets and international benchmarking.

Further to these policy directions, leadership development initiatives have been institutionalised through the Teacher Education Institute of Malaysia (IPGM) and the Institute of Aminuddin Baki (IAB). These institutions have played a central role in professionalising leadership by providing pre-service training, principal preparation programmes and continuous professional development (Ng, 2017, Amirul Fahmie Abdul Razak et al., 2024). According to Adams et al. (2020), Malaysia's investment in leadership preparation reflects an international trend, where the development of leaders is seen as integral to sustaining reform agendas.

3.2. The Current Expectations of School Leaders in Malaysia

During the current reformation age, school leaders in Malaysia are supposed to perform varied roles beyond the normal administration. One of the expectations is accountability. The skills of principals are measured in terms of providing student achievement outcomes, utilizing resources in a transparent and applying ministry instructions (Harris et al., 2017). This focus on accountability is current among the world trends, but in Malaysia it is explicitly high with regard to the centralised policy implementation.

Another expectation is instructional leadership. Siti Noor Ismail et al. (2020) posit that the Malaysian principals are no longer assessed based on their administrative competence but also on their ability to focus on teaching and learning quality. The MEB clearly requests leaders to transform to instructional leaders that facilitate teacher learning, observe classroom practice and curriculum alignment. Meanwhile, community involvement has become important. According to Husaina Banu Kenayathulla et al. (2024), bridge-builders are the new role that effective principals are supposed to play, making alliances with parents, community organisations and external stakeholders to improve the capacity of school. The expectations depict the complex nature of the principalship, which now comprises of accountability, instructional growth and collaboration with the stakeholders.

3.3. The Challenges of Enacting Strategic Leadership

Regardless of the general awareness of the significance of strategic leadership, the school leaders in Malaysia have been struggling with the same problems, which make it hard to implement the concept. These problems can be generally divided into challenges that happen at school level, problems related to teachers, student issues and systemic problems. All of the above categories serve to stress the tensions that leaders need to work with to mobilise sustainable change.

School Challenges

One of the recurring barriers at the school level is the difficulty of shifting from compliance-based, short-term planning to strategies that prioritise long-term sustainability. Leaders often hesitate to take risks and their lack of confidence in initiating change can suppress innovation. As Abdul Rahim Abdul Rahman et al. (2020) point out, poor communication further undermines collaboration by weakening trust and collective purpose. Structural limitations exacerbate these difficulties. Many schools still struggle with inadequate infrastructure, limited funding and outdated technology, conditions that slow the pace of reform. Fred and Bishen Singh (2021) note that these vulnerabilities became especially visible during the COVID-19 pandemic, when schools with weak infrastructure were less able to pivot effectively to online learning. Moreover, inconsistent engagement from parents and communities continues to restrict the ability of leaders to mobilise additional resources, despite the potential of such partnerships to fill important gaps (Nur Annizah Ishak et al., 2020).

Teacher Challenges

Another area of challenge to strategic leadership is represented by teachers. Inexperienced teachers may get overwhelmed in the challenging settings, whereas some of the experienced teachers do not like changes and like usual ways of doing things. The resultant mix of this is lack of flexibility and the inability to implement new initiatives more quickly. Studies indicate that leaders who do not establish enabling, cooperative cultures usually have challenge in eliciting teacher buy-ins on strategies of the school (Mariani Jaafar et al., 2022). Workload is another concern. Extensive administrative demands and excessive documentation requirements frequently divert teachers' attention away from their core instructional responsibilities. The lack of leadership, which actively addresses these pressures, lowers the morale of teachers and decreases classroom performance. Rewards also have a role to play: in the cases where leaders fail to appreciate the contributions of teachers, motivation and professionalism are affected (Jaya & Hidayat, 2020). Moreover, the lack of an equal level of digital literacy in the staff members also limits the effective implementation of technology, an issue that brings to the fore the importance of constant professional growth that should be shaped according to the changing demands of the teachers.

Student Challenges

The students themselves are a complicated issue that determines the way strategic leadership is implemented. Student motivation, attendance and academic achievement are strongly shaped by socio-economic disadvantage. Lower income families tend to have less resources to facilitate learning at home and this has a direct impact on the performance of students. As Rodríguez-Hernández et al. (2020) show, these disparities can significantly undermine school improvement efforts. On top of the socio-economic factor, the Malaysian multicultural environment presents other challenges to school leaders. Language diversity, cultural diversity and value diversity demand that the leaders promote inclusive practices to make all students feel accepted and appreciated. Zatun Najahah Saari and Muhammad Faizal A. Ghani (2024) argue that while diversity enriches school life, it also complicates the task of balancing national education objectives with local realities. For leaders, the challenge lies in fostering inclusive practices that ensure all students feel respected and supported.

Systemic Conditions

Lastly, the presence of systemic conditions is also a major obstacle to strategic leadership. The education system in Malaysia is very centralised with national-level policies and priorities being made. Although this makes it coherent, it limits the ability of school leaders to make reforms that fit their contexts. Principals are usually torn between the accountability requirements by the Ministry and the practical requirements of their schools. As Mestry (2017) as well as Çolak Çakir and Cansoy (2025) observe, this over-centralisation can be counterproductive, placing leaders in a paradoxical position: expected to be visionary and proactive, yet confined by limited decision-making space.

These challenges collectively illustrate that leadership in Malaysian schools is shaped by structural, cultural and systemic constraints that make generic models of leadership insufficient. While international scholarship provides valuable insights, the realities of Malaysia's diverse school communities and centralised governance structure require a framework that is both contextually relevant and practically grounded. It is in response to these conditions that the Strategic Leadership Model for Malaysian School Leaders (STALEM) is derived. The model seeks to translate the demands of policy, practice and research into a coherent set of domains that equip leaders with the capacity to navigate challenges and mobilise sustainable improvement.

4. The Derivation of the Strategic Leadership Model for Malaysian School Leaders (STALEM)

The preceding chapters have examined the theoretical foundations of strategic leadership (Section 2) and the contextual realities of the Malaysian education system (Section 3). Together, these discussions highlight two critical insights. First, strategic leadership provides a powerful lens through which to understand how leaders can guide institutions through uncertainty and reform. Second, leadership models derived from Western contexts, while conceptually robust, may not fully reflect the cultural, social and institutional dynamics of Malaysian schools. These insights underscore the need for a locally relevant framework of strategic leadership that is both theoretically informed and contextually grounded.

The Strategic Leadership Model for Malaysian School Leaders (STALEM) was developed to address this gap. Guided by the principle that leadership frameworks must be responsive to the environments in which they are applied, STALEM integrates global perspectives on strategic leadership with insights derived from Malaysian educational practice. Its development followed a systematic Design and Development Research (DDR) approach, carried out in three phases. The Needs Analysis Phase established the rationale for the model, identifying eight priority needs such as the development of holistic, future-oriented and collaborative leadership guidelines, as well as strategies for addressing school, teacher and student challenges. The Design and Development Phase applied the Modified Delphi Method across three iterative rounds, enabling expert input and refinement of the model's content. This phase produced eight domains, 24 subdomains and 303 items that were validated through consensus among leadership experts. The Evaluation Phase then tested the model's usability using the USE Questionnaire, with findings confirming high levels of acceptance across dimensions of usefulness, ease of use, ease of learning and satisfaction.

The outcome is a comprehensive framework that reflects the strategic responsibilities of Malaysian school leaders. STALEM comprises eight domains: (i) leading change strategies, (ii) leading futuristic thinking skills, (iii) leading vision, mission and values, (iv) leading self-creation, (v) leading stakeholder collaboration, (vi) sustaining a conducive organisational climate, (vii) leading controlled and fair power distribution and (viii) leading strategic competency development programmes. Collectively, these domains articulate a balanced model of leadership that integrates cognitive, behavioural and relational capacities with the unique demands of Malaysian schooling. By situating strategic leadership within the local context, STALEM provides both a theoretical contribution to the literature and a practical guide for enhancing school leadership in Malaysia.

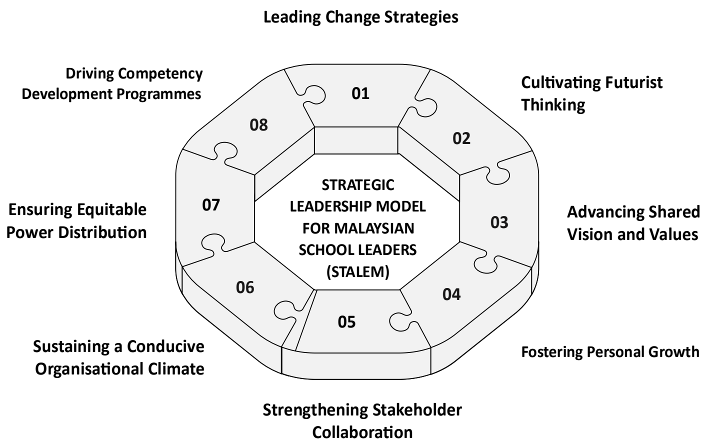

Figure 8. The Strategic Leadership Model for Malaysian School Leaders (STALEM).

Building on this systematic process of derivation, the final version of STALEM consists of eight interrelated domains that collectively capture the strategic leadership responsibilities of Malaysian school leaders. Each domain reflects both global theoretical insights and the specific contextual needs identified through expert consensus and empirical validation. Taken together, they provide a comprehensive framework that guides school leaders in aligning vision with practice, sustaining improvement and responding effectively to the challenges of twenty-first-century education. The following subsections (4.1–4.8) elaborate each of these domains in turn, demonstrating how they function as core dimensions of strategic leadership within Malaysian schools.

4.1. Leading Change Strategies

Successful school transformation demands that the leaders structure intentional and long-term change programs. According to Kotter (1996), urgency creation, coalition building and preservation of gains is the way to attain meaningful change. In Malaysia, where reforms like the Malaysia Education Blueprint are actively guided at the centre, principals have to be flexible to these strategies in order to prevent the superficial approach. By sequencing initiatives systematically and aligning them with school-level priorities, leaders can minimise reform fatigue and foster sustained commitment among teachers and stakeholders. Such an approach enables change to be implemented progressively rather than reactively, increasing the likelihood that improvement efforts are embedded into everyday school practices rather than treated as short-term initiatives.

4.2. Cultivating Futurist Thinking

Strategic leadership involves predictive thinking. According to Slaughter (1998), futurist thinking refers to the capacity to read global trends, predict interruptions and place institutions in sustainable position. The process of futurist thinking in the case of Malaysian schools means equipping the learners with Industry 4.0, digital transformation and global citizenship (Ministry of Education Malaysia, 2013). The principals who practice the foresight strategy of creating scenarios can incorporate the emerging pedagogies and technologies hence future-proofing their learning settings and making learning relevant in a fast-evolving society.

4.3. Advancing Shared Vision and Values

The sense of vision gives unity and values to the team. Bush (2021) stresses that effective communication and articulation of vision can create clarity of direction by leaders and organisational alignment. Leithwood et al. (2020) also say that collective efficacy and resilience are developed using shared values. The inclusivity and cultural legitimacy in the Malaysian educational context are guaranteed by shared vision and value in a multi-ethnic and multi-religious context. Principals foster trust and solidarity that support sustainable improvement, when they align school visions with national objectives and local aspirations of a community.

4.4. Fostering Personal Growth

It is also maintained by strategic leadership whereby leaders themselves promote their learning. Day et al. (2020) point out that good leaders practice reflectively, manage emotions and professional rejuvenation. To Malaysian school leaders, this is represented as enrolling in a formal training through the Institute of Aminuddin Baki (IAB) and the Institutes of Teacher Education (IPGM) as well as informal learning among peers and the community. The modelling of lifelong learning allows principals to sustain the importance of professional development in their respective schools and establish the culture of teacher learning promotion.

4.5. Strengthening Stakeholder Collaboration

School improvement is not autonomous and thus it cannot be sought alone. According to Epstein (2018), good collaboration with parents, communities and other external organisations helps students achieve better and improve schools. In Malaysia, stakeholder cooperation is particularly important because of various cultural and socio-economic diversity of school communities (Abdul Talib Hashim & Amir Hasan Dawi, 2022). By establishing personal contacts with parents, alumni, NGOs, local businesses and universities, strategic leaders increase resources and foster social cohesiveness and build the reputation of schools in the community.

4.6. Sustaining a Conducive Organisational Climate

The success of students and teachers is determined by organisational climate. Al-Zu'bi et al. (2024) highlight that positive climate promotes collaborations, creativity and devotion. To the Malaysian schools, it is possible to maintain such climates by promoting safety, cross-cultural inclusivity and respect and dealing with conflict in constructive ways. By telling the truth, promoting a culture of teamwork and rewarding success, leaders foster the feeling of appreciation and empowerment among teachers, thereby preventing turnover and maintaining the momentum of improvement.

4.7. Ensuring Equitable Power Distribution

The core of collective capacity building is distributed leadership. Spillane (2005) argues that when schools share management of decisions with teachers and middle leaders, they utilise the variety of skills and become more accountable. In Malaysia, equality of power is an alternative to hierarchical traditions but gives an avenue towards common ownership of the reform. Empowerment of teacher committees, student councils and parent groups help the principal decentralise leadership functions so that school improvement is not central to a leader person but cuts across the whole organisation.

4.8. Driving Competency Development Programmes

Sustained improvement relies on teacher capacity. Darling-Hammond (2017) suggest that the quality of instruction and student learning is directly improved by the well-designed professional development. Strategic leaders in Malaysia need to make sure that the competency development programmes are well organized, in line with the national standard and also sensitive to the needs of the school. Teacher actions like coaching, mentoring and professional learning communities (PLCs) enable educators to perfect practice together. Investing in staff capacity, principals make sure that any improvement efforts are converted into literal classroom benefits.

4.9. Integrating the Eight Domains

The discussion of the eight domains illustrates that each represents a distinct yet interconnected aspect of strategic leadership in schools. While individually they address critical areas such as change, vision, collaboration and professional growth, their real strength lies in their integration. STALEM is therefore not intended as a checklist of discrete practices but as a coherent framework in which the domains reinforce one another. Leading change provides momentum, while futurist thinking ensures future readiness. Shared vision and values create coherence, and personal growth sustains leadership capacity. Stakeholder collaboration and organisational climate strengthen relational trust, while equitable power distribution and competency development embed improvement across the school. Together, these pathways demonstrate that strategic leadership in Malaysia is not a collection of isolated practices but a holistic framework that aligns national reforms with local school realities. By operationalising the STALEM model, Malaysian school leaders are better equipped to mobilise communities, foster professional learning and navigate reform challenges. The model thus provides both a conceptual foundation and a practical guide for advancing educational quality and equity.

5. Applications for the STALEM Model

The Strategic Leadership Model for Malaysian School Leaders (STALEM) functions as both a conceptual framework and a practical guide for translating leadership into meaningful school improvement. Its eight interrelated domains provide school leaders with a structured pathway to address complex challenges while remaining aligned with the national education reform agenda. This section discusses how the model can be operationalised in practice, its illustrative applications in school improvement and its integration into leadership preparation and development programmes to ensure long-term sustainability.

5.1. Operationalising the Eight Domains

Kotter (1996) highlights that the change leadership strategies need planned instruments to be used like school improvement plans to bring together short-term priorities and long-term ambitions. Building on this perspective, Slaughter (1998) extends the idea by showing how futurist thinking can be operationalised in terms of scenario planning, horizon scanning and integrating the idea of Education 4.0 into curriculum and pedagogy. In a similar way, Bush (2021) and Leithwood et al. (2020) also indicate that sustaining reform depends on shared vision and values, which are most effective when constructed collaboratively with teachers, students and parents, thereby creating collective ownership and direction.

Other scholars highlight how the eight domains can be operationalised through capacity building and culture. For example, Darling-Hammond et al. (2017) emphasize that personal growth should be promoted with professional development that would be specific to the needs and aspirations of teachers, which would support the culture of lifelong learning. Epstein (2018) shows that enhancing stakeholder partnership relies on long-term partnerships with the parents, alumni and community organisations, which can be used to resource and provide social support beyond the school premises. Similarly, Mousena and Raptis (2020) point out that a favourable school climate emerges from open communication, trust-building practices and reward systems that nurture positive culture.

Leadership practice also involves distributing responsibility and ensuring resilience. Spillane (2005) underscores the value of distributive leadership in balancing power across the organisation, while Lengnick-Hall et al. (2011) argue that long-term competency development builds both staff capacity and institutional resilience. Taken together, these perspectives illustrate how the eight domains of STALEM provide a systematic framework for school leaders to connect vision with practice and to sustain improvement as an ongoing process.

5.2. Illustrative Applications in School Improvement

Day et al. (2020) observe that principals who make use of data-informed change strategies can utilize the student achievement data in order to detect learning gaps, allocate resources efficiently and plan specific professional development. Their writing underscores the importance of instructional supervision that is based on strategic leadership in order to establish conditions in which teaching is constantly observed, assessed and refined. Schools that implement STALEM principles are expected to be in a better position to mobilise parents, alumni and community partners in co-curricular activities, mobilise resources and decision-making processes. This not only enhances resources available in schools but also creates social trust which is a key element in maintaining reform. Spillane (2005) goes further to support this when he demonstrates that with equal power distribution, leadership is no longer the preserve of the principal to middle leaders, teacher leaders and even student councils, but more democratic and participatory schools emerge.

Collectively, these examples indicate how STALEM puts theory to practice to enhance instructional quality, expand stakeholder involvement and develop a culture of shared accountability in schools.

5.3. Integrating STALEM into Leadership Training and Development

Čadová et al. (2024) claim that the training of leadership should not only be aimed at the efficiency in administration but also on the foresight and flexibility. In that regard, organisations like Institute of Aminuddin Baki (IAB) and Teacher Education Institute of Malaysia (IPGM) are in a good position to incorporate STALEM in their curricula so that future and practising leaders can be exposed to the eight areas of strategic leadership. These domains can be taken into practice by case-based learning, simulation of scenarios and structured mentoring. Grimm (2023) and Amirul Fahmie Abdul Razak et al. (2024) are quick to remind us that the process of leadership development should also encompass the idea of professional networks and communities of practice that leaders should share experiences and contextual reviews. Another recommendation of Paletta et al. (2020) that policymakers can also enhance integration is to make school evaluation frameworks consistent with the areas of STALEM to maintain coherence between the leadership practice and accountability. STALEM can be a system of leadership practice, but a system of systemic reform, when built into systems of training and accountability, so that schools can change over time, but improve the quality and equity of education.

6. The Implications for Policy and Practice

As defined and implemented using the STALEM model, strategic leadership has the direct consequences on education reform in Malaysia. Three priorities can be distinguished. To begin with, leadership should be developed by instilling visionary, distributive and adaptive abilities in school leaders, especially by exposing them to trainings programmes in Institute of Aminuddin Baki (IAB) and Teacher Education Institute of Malaysia (IPGM). Second, professional training should not only be based on the initial preparation but lifelong learning possibilities, which are context-specific, like mentoring, peer coaching and networks of colleagues. Third, school governance structures require recalibration to balance centralised policy direction with greater school-level autonomy, enabling schools to innovate and respond to local needs while remaining aligned with national priorities. Collectively, these implications affirm that strategic leadership is a key vehicle for realising both the Malaysia Education Blueprint and Sustainable Development Goals (SDG) 4, ensuring that schools move beyond compliance toward future-ready transformation.

6.1. The Contributions of Strategic Leadership to Malaysia's Reform Agenda

Strategic leadership is directly aligned to the aspirations of Malaysia Education Blueprint by supporting its three overall objectives namely: to increase student outcomes, decrease equity disparities and enhance system efficiency. Through fostering future thinking and a sense of collective vision, school leaders can be placed in a situation to bring the local school interests in line with national interests like Education 4.0 and global competitiveness. Likewise, the focus on stakeholder partnership is also familiar to the Blueprint, which demands a more significant involvement of the community in education. In this regard, STALEM offers a contextually based structure that operationalises school level policy directions that can be used to connect the gap between system ambitions and school realities. Its focus on inclusion, equity and sustainability also resonates strongly with SDG 4, reinforcing the principle that quality education must extend to all learners regardless of background.

6.2. The Recommendations for Leadership Development, Professional Training and Governance

Implementation of strategic leadership needs to be long-term leadership development and governance reform. Three priorities emerge:

1. Leadership Development: Leadership preparation must be transforming the eight domains of STALEM into training modules with the focus on visionary thinking, distributive practices and ability

to lead change. It can be done by case-based learning, simulations and reflective practice that are within programmes conducted by IAB and IPGM. This contributes to SDG 4 by ensuring leaders are equipped to promote inclusive and equitable learning environments.

2. Professional Training: Professional training ought not to be only the initial preparation, it must also extend to in-service training. The training opportunities should be dynamic, individualised and connected to the actual school issues. Mentoring, peer coaching and professional networks mechanisms may be platforms where STALEM principles may be transformed into practice. This also SDG 4's emphasis on lifelong learning by creating systems where teachers and leaders are continually learning and adapting.
3. School Governance: The governance structures should establish facilitative conditions of strategic leadership. This involves decentralising some of the decision-making authority to schools with accountability ensured by open evaluation systems. There should be a balance between a centralised policy direction and school-based autonomy where the leaders should be able to localise reforms whenever it fits their context without going off track with national interests.

6.3. The Relevance for Policymakers, Training Institutes, and School Leaders

The policymakers should incorporate strategic leadership in more pervasive reforms so that there is consistency between the national interests and school-based practices (Mariani Jaafar et al., 2022). In the case of training institutes like IAB and IPGM, the focus is on preventing the movement of programme design to involve a limited administrative competence, but the development of strategic capacity and foresight and resilience. Meanwhile, school leaders are also advised to incorporate STALEM as a guide and a reflection tool. Through the assessment of progress in its eight domains, the principals are able to determine areas of growth, marshal support and hold vision and practice to account.

Taken together, these implications reaffirm that strategic leadership is one of the key mechanisms of promoting education reform agenda in Malaysia. By connecting national policies with the global aspirations of SDG 4, STALEM provides a pathway for schools to transition from compliance-driven practices to future-oriented, innovative, and inclusive environments.

Conclusion

This chapter has looked at how strategic leadership can help schools in Malaysia improve in a way that lasts. It followed the development of leadership discourse from the National Education Philosophy to the Malaysia Education Blueprint, showing how school leaders are being asked to be responsible administrators, instructional leaders and community bridge-builders more and more. The investigation highlighted ongoing difficulties such as limited resources, teacher preparedness, student diversity and institutional centralization, that continue to influence leadership practices in Malaysian schools. The chapter introduced the Strategic Leadership Model for Malaysian School Leaders (STALEM) in response to these facts. Unlike other models of leadership, STALEM is created for Malaysia and includes eight related areas that include both global issues and national goals. The approach offers a useful framework for school leaders to deal with structural problems, get everyone to work together and work for long-term change.

This work is novel because it provides a model that is based on real-life situations and brings together global ideas about strategic leadership with the unique aspects of Malaysia's education system. In this way, it shows how important it is to have frameworks that are both theoretically sound and useful in real life. Subsequent research may further evaluate and enhance STALEM in various educational environments, investigating its relevance across different educational stages and its influence on student performance. The model shows policymakers that school leaders require more freedom, professional development and resources to be able to do strategic leadership well. To move forward with strategic leadership in Malaysia, school leaders need to be able to balance their big dreams with the realities of their situations. This balance is essential to ensuring that school leaders are able to enact meaningful and sustainable change.

References

Abdul Rahim Abdul Rahman, Lokman Mohd Tahir, Siti Nisrin Mohd Anis, & Mohd Fadzli Ali. (2020). Exploring challenges in practicing instructional leadership: Insights from senior secondary principals. *Universal Journal of Educational Research, 8*(11), 83–96. https://doi.org/10.13189/ujer.2020.082310

Abdul Talib Hashim, & Amir Hasan Dawi. (2022). A case study of racial diversity in a multicultural school environment. *International Journal of Academic Research in Business and Social Sciences, 12*(10), 691–701. https://doi.org/10.6007/IJARBSS/v12-i10/14615

Adair, J. E. (2002). *Effective strategic leadership*. Macmillan.

Adair, J. E. (2005). *How to grow leaders: The seven key principles of effective leadership development*. Kogan Page.

Adams, D., Mooi, A. N. Y., & Muniandy, V. (2020). Principal leadership preparation towards high-performing school leadership in Malaysia. *Asian Education and Development Studies, 9*(4), 425–439. https://doi.org/10.1108/AEDS-02-2018-0046

Al-Zu'bi, M., Al-Mseidin, K., Yaakob, M., Fauzee, M., Mahmood, M.-H.-H., Sohri, N., Al-Mawadieh, R., & Senathirajah, A. R. (2024). The impact of the school educational environment on primary school teachers' commitment to educational quality. *Discover Sustainability*. https://doi.org/10.1007/s43621-024-00633-4

Al Huraizi, S., & Marni, N. (2023). Strategic leadership practices and their role in achieving competitive advantage in schools: Future agendas, research gaps, and research needs. *International Journal of Academic Research in Business and Social Sciences, 13*(2), 819–832. https://doi.org/10.6007/IJARBSS/v13-i2/16424

Alan, H., & John, P. (2001). *Intelligent leadership: Creating a passion for change*. Random House Business Books.

Amels, J., Krüger, M., Suhre, C., & Van Veen, K. (2021). The relationship between primary school leaders' utilisation of distributed leadership and teachers' capacity to change. *Educational Management Administration & Leadership, 49*(5), 732–749. https://doi.org/10.1177/1741143220915921

Amirul Fahmie Abdul Razak, Muhammad Faizal A. Ghani, & Norfariza Mohd Radzi. (2024). Amalan kepimpinan guru pelatih Institut Pendidikan Guru Malaysia di Zon Tengah. *Jurnal Kepimpinan Pendidikan (JuPiDi), 11*(2), 35–51. https://doi.org/10.22452/jupidi.vol11no2.3

Amirul Fahmie Abdul Razak, Muhammad Faizal A. Ghani, Norfariza Mohd Radzi, & Nur Diyana Zakariah. (2024). Developing leadership for pre-service teachers: Previous concepts and models. *Akademika, 94*(3), 181–203.

Bass, B. M. (2007). Executive and strategic leadership. *International Journal of Business, 12*(1), 33–52.

Beare, H. (2013). *Creating the future school*. Routledge.

Boal, K. B., & Bryson, J. M. (1988). Charismatic leadership: A phenomenological and structural approach. *Academy of Management Review, 13*(1), 11–28.

Boal, K. B., & Hooijberg, R. (2001). Strategic leadership research: Moving on. *The Leadership Quarterly, 11*(4), 515–549. https://doi.org/10.1016/S1048-9843(00)00057-6

Bush, T. (2020). *Theories of educational leadership and management* (5th ed.). Sage.

Bush, T. (2021). Assessing successful school leadership: What do we know? *Educational Management Administration & Leadership, 49*(5), 687–689. https://doi.org/10.1177/17411432211034675

Čadová, L., Stříbrný, J., & Sládek, P. (2024). Leadership in the digital age: Navigating complexity through critical thinking, creativity, and unconventional approaches. *Journal of Interdisciplinary Research, 14*(1), 234–239. https://doi.org/10.33543/j.1401.234239

Chandler, A. D. (1962). *Strategy and structure: Chapters in the history of the industrial enterprise*. MIT Press.

Cheng, Y. C. (2019). *Paradigm shift in education: Towards the third wave of effectiveness*. Routledge.

Çolak Çakir, T., & Cansoy, R. (2025). Challenges of being a principal in vocational high schools perceived as "a last chance": An examination in a centralized education system. *Leadership and Policy in Schools*. https://doi.org/10.1080/15700763.2025.2493727

Cowley, W. H. (1928). Three distinctions in the study of leaders. *The Journal of Abnormal and Social Psychology, 23*(2), 144–157. https://doi.org/10.1037/h0073661

Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education, 40*(3), 291–309. https://doi.org/10.1080/02619768.2017.1315399

Darling-Hammond, L., Burns, D., Campbell, C., Goodwin, A. L., Hammerness, K., Low, E.-L., McIntyre, A., Sato, M., & Zeichner, K. (2017). *Empowered educators: How high-performing systems shape teaching quality around the world*. John Wiley & Sons.

Davies, B. (2009). *The essentials of school leadership*. Sage.

Davies, B. J., & Davies, B. (2004). Strategic leadership. *School Leadership & Management, 24*(1), 29–38. https://doi.org/10.1080/1363243042000172804

Day, C., Gu, Q., & Sammons, P. (2016). The impact of leadership on student outcomes: How successful school leaders use transformational and instructional strategies to make a difference. *Educational Administration Quarterly, 52*(2), 221–258.

Day, C., Sammons, P., & Gorgen, K. (2020). *Successful school leadership*. Routledge.

Drucker, P. F. (1996). Your leadership is unique. *Leadership, 17*(4), 54.

Eacott, S. (2008a). An analysis of contemporary literature on strategy in education. *International Journal of Leadership in Education, 11*(3), 257–280. https://doi.org/10.1080/13603120701462111

Eacott, S. (2008b). Strategy in educational leadership: In search of unity. *Journal of Educational Administration, 46*(3), 353–375. https://doi.org/10.1108/09578230810869284

Eacott, S. (2011). Leadership strategies: Re-conceptualising strategy for educational leadership. *School Leadership & Management, 31*(1), 35–46. https://doi.org/10.1080/13632434.2010.540559

Epstein, J. L. (2018). *School, family, and community partnerships: Preparing educators and improving schools* (Student economy ed.). Routledge.

Fazleen Mohamad, & Siti Noor Ismail. (2018). Kepimpinan strategik dan hubungannya dengan kualiti pengajaran guru. *JuPiDi: Jurnal Kepimpinan Pendidikan, 5*(2), 14–25.

Fiedler, F. E. (1967). *A theory of leadership effectiveness*. McGraw-Hill.

Finkelstein, S., Hambrick, D. C., & Cannella, A. A. (2009). *Strategic leadership: Theory and research on executives, top management teams, and boards*. Oxford University Press.

Fred, A., & Bishen Singh, G. S. (2021). Instructional leadership practices in under-enrolled rural schools in Miri, Sarawak. *Asian Journal of University Education (AJUE), 17*(1), 165–176. https://doi.org/10.24191/ajue.v17i1.12694

Fullan, M. (2004). *Leadership & sustainability: System thinkers in action*. Corwin Press.

Fullan, M. (2014). *The principal: Three keys to maximizing impact*. Jossey-Bass.

Funda, K. (2022). The role of strategic leadership in innovation performance. *Open Journal of Business and Management, 10*, 654–669. https://doi.org/10.4236/ojbm.2022.102037

Gardner, H. (2011). *Leading minds: An anatomy of leadership*. Basic Books.

Gill, R. (2013). *Theory and practice of leadership* (2nd ed.). Sage.

Gioia, D. A., & Thomas, J. B. (1996). Identity, image, and issue interpretation: Sensemaking during strategic change in academia. *Administrative Science Quarterly, 41*(3), 370–403.

Grimm, F. (2023). Teacher leadership for teaching improvement in professional learning communities. *Professional Development in Education, 50*(6), 1135–1147. https://doi.org/10.1080/19415257.2023.2264286

Hallinger, P. (2011). Leadership for learning: Lessons from 40 years of empirical research. *Journal of Educational Administration, 49*(2), 125–142.

Hambrick, D. C. (1989). Guest editor's introduction: Putting top managers back in the strategy picture. *Strategic Management Journal, 10*(S1), 5–15.

Hambrick, D. C., & Fredrickson, J. W. (2005). Are you sure you have a strategy? *Academy of Management Perspectives, 19*(4), 51–62. https://doi.org/10.5465/ame.2005.19417907

Harris, A., Jones, M., Cheah, K., Devadason, E., & Adams, D. (2017). Exploring principals' instructional leadership practices in Malaysia: Insights and implications. *Journal of Educational Administration, 55*(2), 207–221. https://doi.org/10.1108/JEA-05-2016-0051

Hopkins, D. (2001). *School improvement for real*. Routledge. https://doi.org/10.4324/9780203165799

Husaina Banu Kenayathulla, Muhammad Faizal A. Ghani, & Norfariza Mohd Radzi. (2024). Enhancing quality appointment, preparation, and support systems for Malaysian principals. *Education Sciences, 14*(6), Article 659. https://doi.org/10.3390/educsci14060659

Ibrahim Mamat. (1998). *Pengetua sekolah menangani isu dan cabaran kepemimpinan*. Kumpulan Budiman.

Ireland, R. D., & Hitt, M. A. (1999). Achieving and maintaining strategic competitiveness in the 21st century: The role of strategic leadership. *Academy of Management Perspectives, 13*(1), 43–57. https://doi.org/10.5465/ame.1999.1567311

Jabbar, A., & Hussein, A. (2017). The role of leadership in strategic management. *International Journal of Research – Granthaalayah, 5*(5), 99–106. https://doi.org/10.29121/granthaalayah.v5.i5.2017.1841

Jäppinen, A.-K. (2017). Analysis of leadership dynamics in educational settings during times of external and internal change. *Educational Research, 59*(4), 460–477. https://doi.org/10.1080/00131881.2017.1376591

Jaya, W. S., & Hidayat, N. (2020). The relationship of school leadership and teacher attitudes towards professional competency. *International Journal of Multicultural and Multireligious Understanding, 7*(10), 503–509. https://doi.org/10.18415/ijmmu.v7i10.2107

Joyce, P. (2012). *Strategic leadership in the public services*. Routledge. https://doi.org/10.4324/9780203608937

Kabeyi, M. (2019). Organizational strategic planning, implementation, and evaluation: Challenges and benefits. *International Journal of Applied Research, 5*(6), 27–32.

Ko, J. Y. C., Hallinger, P., & Walker, A. D. (2012). Exploring school improvement in Hong Kong secondary schools. *Peabody Journal of Education, 87*(2), 216–234.

Kotter, J. P. (1996). *Leading change*. Harvard Business School Press.

Leithwood, K., Harris, A., & Hopkins, D. (2020). Seven strong claims about successful school leadership revisited. *School Leadership & Management, 40*(1), 5–22.

Lengnick-Hall, C., Beck, T., & Lengnick-Hall, M. (2011). Developing a capacity for organizational resilience through strategic human resource management. *Human Resource Management Review, 21*(3), 243–255. https://doi.org/10.1016/j.hrmr.2010.07.001

Luciano, M. M., Nahrgang, J. D., & Shropshire, C. (2020). Strategic leadership systems: Viewing top management teams and boards of directors from a multiteam systems perspective. *Academy of Management Review, 45*(3), 675–701. https://doi.org/10.5465/AMR.2017.0485

Mariani Jaafar, Soaib Asimiran, Arnida Abdullah, & Siti Noormi Alias. (2022). Strategic leadership practices among school leaders in Malaysia: A systematic literature review. *International Journal of Academic Research in Business and Social Sciences, 12*(3), 1–18. https://doi.org/10.6007/IJARBSS/v12-i3/12192

Mestry, R. (2017). Empowering principals to lead and manage public schools effectively in the 21st century. *South African Journal of Education, 37*(1), Article a1334. https://doi.org/10.15700/saje.v37n1a1334

Mincu, M. (2022). Why is school leadership key to transforming education? Structural and cultural assumptions for quality education in diverse contexts. *Prospects, 52*, 231–242. https://doi.org/10.1007/s11125-022-09625-6

Ministry of Education Malaysia. (1988). *National Philosophy of Education (Falsafah Pendidikan Kebangsaan)*. Ministry of Education Malaysia.

Ministry of Education Malaysia. (2013). *Malaysia Education Blueprint 2013–2025*. Ministry of Education Malaysia.

Mintzberg, H. (1994). Rethinking strategic planning part I: Pitfalls and fallacies. *Long Range Planning, 27*(3), 12–21. https://doi.org/10.1016/0024-6301(94)90185-6

Mousena, E., & Raptis, N. (2020). Beyond teaching: School climate and communication in the educational context. In *Education at the intersection of globalization and technology*. IntechOpen. https://doi.org/10.5772/intechopen.93575

Muhammad Faizal A. Ghani. (2013). Development of an effective school model for Malaysian schools. *International Journal of Academic Research, 5*(5), 131–142. https://doi.org/10.5829/idosi.mejsr.2014.19.10.11483

Neyişci, N., & Erdoğan, A. (2022). Strategic leadership and its reflections in educational organizations. *Journal for Educators, Teachers and Trainers, 13*(3), 15–22. https://doi.org/10.47750/jett.2022.13.03.002

Ng, A. (2017). School leadership preparation in Malaysia. *Educational Management Administration & Leadership, 45*(6), 1002–1019. https://doi.org/10.1177/1741143216662922

Northouse, P. G. (2021). *Leadership: Theory and practice* (9th ed.). Sage.

Nur Annizah Ishak, Nurul Huda Mohd Satar, & Roza Hazli Zakaria. (2020). Parental involvement in education among urban families in Malaysia. *JATI: Journal of Southeast Asian Studies, 25*(2), 60–85. https://doi.org/10.22452/jati.vol25no2.4

Nur Hidayah, P., Ernie Tisnawati, S., Yuyun, W., & Septiadi, P. (2015). How to develop strategic leadership in higher education institutions. *International Journal of Economics, Commerce and Management, 3*(5), 1164–1175.

Paletta, A., Ferrari, E. B., & Alimehmeti, G. (2020). How principals use a new accountability system to promote change in teacher practices: Evidence from Italy. *Educational Administration Quarterly, 56*(1), 123–173. https://doi.org/10.1177/0013161X19840398

Perez, D., & Banayo, A. (2023). Leadership practices and management behavior of school heads towards quality service and teacher performance. *International Journal of Research Publications*. https://doi.org/10.47119/ijrp1001271620235061

Pisapia, J. (2009). *The strategic leader: New tactics for a globalizing world*. Information Age Publishing.

Prahalad, C. K., & Hamel, G. (2009). The core competence of the corporation. In *Knowledge and strategy* (pp. 41–59). Routledge.

Pryor, M. G., Anderson, D., Toombs, L. A., & Humphreys, J. H. (2007). Strategic implementation as a core competency: The 5P's model. *Journal of Management Research, 7*(1), 3–17.

Rodríguez-Hernández, C. F., Cascallar, E., & Kyndt, E. (2020). Socio-economic status and academic performance in higher education: A systematic review. *Educational Research Review, 29*, Article 100305. https://doi.org/10.1016/j.edurev.2019.100305

Selvaratnam, V. (2022). Education: The past and the future. In *Malaysia's leap into the future: The building blocks towards balanced development* (pp. 183–213). Springer.

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. Doubleday.

Siti Noor Ismail, Shamsuddin Muhammad, Mohd Norakmar Omar, & Arumugam Raman. (2020). The great challenge of Malaysian school leaders' instructional leadership: Can it affect teachers' functional competency across 21st-century education? *Universal Journal of Educational Research*.

Slaughter, R. A. (1998). Futures beyond dystopia. *Futures, 30*(10), 993–1002. https://doi.org/10.1016/S0016-3287(98)00101-3

Spillane, J. P. (2005). Distributed leadership. *The Educational Forum, 69*(2), 143–150.

Stacey, R. D. (1992). *Managing the unknowable: Strategic boundaries between order and chaos in organizations*. Wiley.

Stogdill, R. M. (1974). *Handbook of leadership: A survey of theory and research*. Free Press.

Suharto, S. (2023). Strategic leadership: Effective leadership in strategic management. *Jurnal Ilmiah Multidisiplin Indonesia (JIM-ID), 2*(1), 38–44. https://doi.org/10.58471/esaprom.v2i01.3959

Tichy, N. W., & Sherman, S. (1993). *Control your destiny or someone else will: How Jack Welch is turning General Electric into the world's most competitive corporation*. Doubleday.

Wasserberg, M. (1999). Creating the vision and making it happen. In *Living headship: Voices, values and vision* (p. 208). Sage.

Williams, H., & Johnson, T. (2013). Strategic leadership in schools. *Education 3–13, 41*(3), 350–355.

Yukl, G. (2012). *Leadership: Theory and practice* (7th ed.). Sage.

Zatun Najahah Saari, & Muhammad Faizal A. Ghani. (2024). Development of cross-cultural leadership model (m2cL) for Malaysian school leaders. *Management Research Journal, 13*(1), 20–34. https://doi.org/10.37134/mrj.vol13.1.2.2024

Chapter 4

An At-Home Acid–Base Titration of Vinegar Using a POGIL-Based Inquiry Framework

Sevinç Nihal Yeşiloğlu*
and Yüksel Altun
Division of Chemistry Education, Gazi Faculty of Education, Gazi University, Ankara, Türkiye

Abstract

This study describes a safe and at-home acid–base titration activity using household materials, designed within a Process-Oriented Guided Inquiry Learning (POGIL) framework to support conceptual rigor during remote chemistry instruction. Nineteen second-year undergraduate pre-service chemistry teachers titrated diluted vinegar with sodium bicarbonate and used red cabbage extract as a natural pH indicator. The activity was conducted individually in students' home kitchens environments and supported by structured online group discussions, with an emphasis on experimental reasoning and conceptual understanding. Students identified titration endpoints through observable color changes and gas evolution, applying Le Châtelier's principle to explain the role of CO_2 removal. Despite the use of non-standardized equipment, the measured acetic acid concentrations ranged from 2.6% to 5.1% (w/v), with a class mean of 3.8% (w/v) and a standard deviation of 0.83%. The findings demonstrate that POGIL-based, home-adapted experiments can support meaningful scientific learning. The approach also enabled exploration of equilibrium, indicators, and experimental design, offering a replicable model for remote and resource-limited science education settings.

* Corresponding Author's Email: nihalatalay@gazi.edu.tr

In: Progress in Education. Volume 91
Editor: Roberta V. Nata
ISBN: 979-8-90134-375-3

Keywords: acid-base titration; POGIL; at-home laboratory

Introduction

Acid–base titration is a fundamental analytical technique in chemistry education used to determine the concentration of an acid or base in a solution. In a typical laboratory setting, a strong base or a strong acid titrant is used to titrate a weak acid or a weak base, ensuring the neutralization reaction completes and yields a clear equivalence point. The reason a strong titrant is preferred is that a weak acid titrated with a weak base generally does not have a sharp endpoint; the reaction may not proceed to completion, and the pH change is more gradual, making the endpoint difficult to detect (Chang & Goldsby, 2021). During the COVID-19 pandemic, however, many instructors sought innovative approaches to enable students to perform titrations at home using readily available materials (Andrews et al., 2020). White vinegar (which contains acetic acid) and baking soda (sodium bicarbonate) are common household chemicals that can, in principle, be used to perform an acid–base titration. A key limitation is that sodium bicarbonate is a weak base; simply mixing vinegar and baking soda in a closed system would result in an incomplete neutralization reaction due to the equilibrium between acetic acid and bicarbonate.

However, if one of the reaction products of the neutralization reaction is continuously removed, the equilibrium can be shifted toward completion according to Le Châtelier's principle (Atkins & de Paula, 2022; Chang & Goldsby, 2021). In the case of acetic acid (CH_3COOH) and bicarbonate (HCO_3^-), the neutralization produces carbonic acid (H_2CO_3), which then decomposes into carbon dioxide gas (CO_2) and water (see Equation 1). The evolving CO_2 escapes the solution, effectively removing a reaction product and shifting the equilibrium toward completion of the neutralization. The net reaction is thus:

$$CH_3COOH(aq) + NaHCO_3(aq) \rightleftharpoons CH_3COONa(aq) + CO_2(g) + H_2O(l) \quad (1)$$

Despite the beneficial removal of CO_2, one must be cautious: as CO_2 dissolves in water, it forms carbonic acid, which is weakly acidic. If CO_2 is not removed from the solution, the extra carbonic acid can cause the solution's pH to drop again, obscuring the titration endpoint. Therefore, effective removal of CO_2 from the solution is essential for accurate endpoint determination. In a laboratory, one might heat the reaction flask to reduce CO_2 solubility or use stirring to encourage outgassing. Henry's law tells us that the solubility of a gas in a liquid is directly proportional to the gas's partial pressure above the liquid, and solubility is highly temperature-dependent (Harrison & De Mora, 1996). As temperature increases, the solubility of gases decreases. Thus, heating the reaction mixture can significantly accelerate the evolution CO_2 by reducing gas solubility. In this at-home titration, instead of using a hot plate—which is typically unavailable in most students' home kitchens—the reaction vessel was placed in near-boiling water to maintain an elevated temperature (approximately 75–80 °C), thereby facilitating the continuous removal of CO_2. The necessity of heating is addressed later, examining whether it is strictly required or accelerates the reaction. Qualitatively, CO_2 would eventually escape without heating—allowing the reaction to reach completion given sufficient time and agitation—but heating ensures rapid gas evolution, enabling the endpoint to be observed within a limited laboratory period.

The present study draws upon the work of Andrews et al. (2020), who provided detailed descriptions of the substances and materials used by students to conduct acid–base titrations at home, along with the corresponding procedures. Guided by the insights gained, the researchers developed the activity sheets within the established framework of Process-Oriented Guided Inquiry Learning (POGIL). This approach was intentionally designed to structure the learning experience in a manner that promotes epistemic engagement and conceptual understanding. The rationale for employing the POGIL approach in the activity design is further supported by recent findings in the literature regarding students' conceptual challenges in learning acid–base titrations. Salame and Montero (2022) surveyed students' understanding of acid–base titration concepts and found that many students have difficulties connecting the macroscopic titration procedure with the underlying microscopic acid–base equilibria. They reported an over-reliance on rote procedures and calculations, rather than conceptual understanding, when students analyze titration problems. Such findings underscore the need for instructional approaches that promote deeper conceptual engagement. The activity in this study addresses this issue by using a POGIL framework, which encourages students to construct understanding through targeted questions and group discussion, rather than following a cookbook procedure.

Numerous instructors devised at-home laboratory exercises during the pandemic to replace traditional in-person laboratories. Several at-home acid–

base titration activities have been reported in the literature. For instance, Andrews et al. (2020) implemented a four-week at-home general chemistry lab program at the University of Southern California during the COVID-19 pandemic. The experiments, which included acid–base titrations among other topics, were designed to use only safe household materials and required no specialized equipment. Caraballo et al. (2021) developed a set of low-cost, environmentally friendly acid–base equilibrium experiments for at-home use, particularly suited to remote teaching contexts. These activities employed household materials and natural indicators such as curcumin from turmeric and red cabbage, with digital RGB image analysis via smartphone to effectively monitor pH changes and titration endpoints. Nguyen et al. (2023) introduced a "lemon juice titration" in which citric acid in lemon juice was titrated with baking soda solution at home; notably, they used Mentos candies to nucleate CO_2 bubbles and hasten outgassing. This approach allowed a weak acid (citric)–weak base (bicarbonate) titration to reach an endpoint by using vigorous gas evolution as an indicator. Their students' at-home results were comparable to traditional in-lab titrations with NaOH, highlighting that hands-on inquiry can be successfully translated to the home environment. Another recent innovation is the use of smartphone technology. Li et al. (2023) developed a mobile app for quantifying pH using test strips, allowing students to monitor titrations more quantitatively at home by reading subtle pH changes through the phone's camera. Such technology-enhanced approaches offer increased precision, although they require access to compatible devices and proper calibration routines.

Compared to prior studies, the present work contributes by integrating a home-adapted titration experiment with a structured POGIL-based pedagogical framework. The students in the study were not merely following instructions to do a titration; instead, the instructor guided them through a structured sequence of questions that prompted them to predict, observe, explain, and reflect on the experimental process. For example, students had to decide when they had reached the titration endpoint based on qualitative observations rather than being told a specific volume in advance. They also had to determine how to measure and combine materials using nonstandard apparatus (e.g., using a teaspoon as a measure of solid baking soda, counting drops from a syringe as a measure of liquid) and consider how experimental conditions such as using hot vs. cold water would affect the reaction. Through these decisions, students engaged in authentic scientific reasoning, fostering deeper conceptual understanding of acid–base titration and equilibrium processes.

Process-Oriented Guided Inquiry Learning (POGIL)

Process-Oriented Guided Inquiry Learning (POGIL) is a student-centered instructional approach that integrates guided inquiry with developing process skills such as teamwork, communication, and critical thinking. Unlike traditional lecture-based instruction, POGIL engages students as active participants in constructing their understanding through structured, collaborative learning activities. In a typical POGIL classroom, students work in small self-managed teams on specially designed materials that guide them to explore concepts, invent relationships, and apply their understanding to new contexts.

POGIL originated in the 1990s as part of a National Science Foundation initiative aimed at improving the effectiveness of chemistry instruction at the collegiate level (Farrell et al., 1999). It was first implemented in general chemistry courses, where educators observed that many students struggled with conceptual understanding despite success in algorithmic problem solving. In response, instructors developed inquiry-based learning activities promoting conceptual development while fostering important scientific process skills (Moog & Spencer, 2008).

Extensive research has since demonstrated the effectiveness of POGIL in improving student learning outcomes in chemistry. For example, students in POGIL-based general chemistry courses have shown increased conceptual understanding, improved retention rates, and enhanced problem-solving abilities compared to traditional lecture courses (Lewis & Lewis, 2005; Ruder & Hunnicutt, 2008). Moreover, the structured collaborative format of POGIL provides a framework for meaningful engagement in laboratory environments. In chemistry labs, POGIL has been adapted to replace step-by-step instructions with guided-inquiry experiments that prompt students to predict, observe, and explain chemical phenomena (Bretz, 2013; Hanson, 2006). Such implementations promote a more profound understanding of experimental design and reinforce the connections between theoretical knowledge and laboratory practice.

The POGIL-PCL (Process-Oriented Guided Inquiry Learning–Physical Chemistry Laboratory) framework, developed by Hunnicutt et al. (2015), offers a structured model for designing laboratory experiments that engage students in multiple “data-thinking cycles.” In this model, students begin with a conceptual question that frames the investigation and proceed through iterative cycles of predicting, observing, and reasoning. Each cycle is designed to deepen conceptual understanding by requiring students to generate hypotheses, follow a standard protocol, and analyze collected data, often in collaboration with peers. The experiments are intentionally structured to require at least two such cycles—moving from qualitative observations toward quantitative modeling—thereby promoting deeper engagement with underlying chemical principles.

Informed by this framework, this study adapted the POGIL-PCL model to a single-cycle structure consisting of four key stages: (1) pre-experiment questions, (2) prediction and observation tasks, (3) thinking-about-data questions, and (4) post-experiment reflection. This adaptation preserved the core features of guided inquiry while streamlining the process to suit the constraints of an at-home experimental setting. By maintaining an emphasis on prediction, real-time observation, and data-based reasoning, the activity aimed to foster meaningful conceptual engagement within a more focused and accessible inquiry cycle. This process also parallels the stages of Kolb's experiential learning cycle—concrete experience, reflective observation, abstract conceptualization, and active experimentation (Kolb et al., 2014) by encouraging students to construct meaning from their observations and apply chemical concepts within an experimental context. Through this integration of inquiry and experience, the activity supported both individual sensemaking and collaborative conceptual development. Furthermore, using POGIL in small teams fosters socially constructed understanding, consistent with Vygotsky's (1978) view of learning as a socially mediated process.

To delineate the scope of this chapter, the study was guided by the following research question: To what extent is a POGIL-based at-home titration activity effective in enhancing undergraduate students' conceptual understanding of acid–base equilibrium through common household materials? In addressing this question, the investigation sought to assess the pedagogical impact of the activity—namely, student engagement and the development of process skills—and its effectiveness in fostering conceptual understanding of titration chemistry and equilibrium principles within a guided-inquiry framework implemented in a home setting.

The remainder of this chapter is structured as follows. First, the experimental design and implementation of the at-home titration activity are described, including details of the participants, materials used, procedure followed, and the questions on the POGIL worksheet. This is followed by presenting results and discussion, encompassing typical student outcomes, evidence of student learning derived from written responses and group discussions, and an analysis of how the inquiry-based format influenced their conceptual understanding. Specific issues are also addressed, such as the role of heating in driving the equilibrium to completion and the alignment of the activity with learning objectives including applications of Le Châtelier's principle, pH indicators, and experimental uncertainty. An analysis of representative student responses and a critique of common misconceptions observed are provided. The chapter concludes with a summary of student learning outcomes, practical recommendations for improving the activity, and suggestions for future research and broader application, such as incorporating this at-home experiment into teacher education workshops or comparing its efficacy to traditional lab approaches.

Methodology

Participants and Context

The activity was conducted during the COVID-19 pandemic within the Analytical Chemistry Laboratory II course at a public university in Türkiye, involving 19 second-year undergraduate pre-service chemistry teachers (16 females, three males, aged 19–22).

This laboratory course was delivered remotely via a Learning Management System (LMS) supported by a synchronous virtual conferencing platform. The platform enabled real-time interaction, including audio–visual communication, screen sharing, and collaborative discussion tools. The activity was implemented over four weeks of scheduled online sessions, each lasting approximately 1.5 hours.

In the first online session, students were introduced to the fundamental principles of acid-base titration. They watched brief instructional videos on titration technique and discussed key concepts such as indicators, the endpoint vs. the equivalence point, and the completeness of reactions. Following this introduction, students were briefed on the details of the upcoming at-home activity. They were assigned to small groups of 3–4 members (POGIL teams) to work together and were encouraged to communicate and collaborate outside of scheduled class if needed. Prior to the experiment, students received preparatory study questions as pre-laboratory preparation to ensure readiness for the relevant concepts (described below). They were also given a detailed list of the materials and reagents they would need to obtain or prepare in advance. Each student acquired the required household materials and conducted the titration independently in their home environment, typically in a kitchen setting. Students remained in active interaction with peers and instructors during the scheduled online sessions. During the second online session (one week before the experiment), the instructor provided feedback on the preparatory questions. The third session was the implementation of the titration activity itself, conducted synchronously with all students performing the experiment in parallel while in the video meeting. The final session was dedicated to discussing the experimental results, analyzing the data, post-experiment questions, and reflecting on the activity.

Learning Objectives

The learning objectives for this activity were designed to address content knowledge and scientific process skills. In terms of content, students were expected to:

- Develop an understanding of the role of pH indicators in acid–base titrations, and distinguish between the endpoint, equivalence point, and half-equivalence point in the titration process.
- Determine the titration endpoint using an appropriate indicator and calculate the acidity (molarity and percent by mass of acetic acid) of a vinegar sample based on titration data.
- Describe the chemical changes occurring during a weak acid–weak base titration by writing balanced chemical equations and considering the relevant equilibria (including the formation and removal of carbonic acid and CO_2).
- Construct or interpret a titration curve for a weak acid titrated with a weak base, reinforcing their understanding of buffer regions, equivalence point pH, and titration theory.

In addition to content-based objectives, the activity aimed to develop students' scientific process skills and inquiry abilities. Students were expected to observe and interpret color changes resulting from pH variations, record qualitative and quantitative observations diligently, analyze experimental data accurately, and connect their empirical findings with theoretical concepts to draw scientifically sound conclusions. The collaborative POGIL structure also targeted process skills such as teamwork, communication, and self-directed learning.

Materials

All materials were selected based on accessibility, safety, and pedagogical relevance, and were readily obtainable in a household context. Each student was instructed to gather the following in advance:

- White vinegar: a standard grocery product containing acetic acid. White vinegar served as the acid analyte.
- Baking soda (sodium bicarbonate, $NaHCO_3$): a common household base, used as the titrant.
- Natural pH indicator (red cabbage extract): prepared by each student beforehand by boiling red cabbage leaves in water. Approximately 2 cups (approximately 500 mL) of red cabbage solution were prepared, yielding a purple extract. (Instructor note: Students were told this indicator solution can be stored in a refrigerator for up to 3 days without spoilage.)
- Hot water: near-boiling water (approximately 75–80 °C), obtained by boiling water in a kettle or pot. This hot water was used as a heat bath to keep the reaction vessel warm during titration, reducing CO_2 solubility as described.
- Measuring implements: standard kitchen measuring tools were used instead of laboratory glassware. For example, a tablespoon (approximately 10-15 mL, depending on calibration) and a teaspoon (approximately 5 mL) were used to measure liquids and solids. Based on prior calibration by the instructors, one heaped teaspoon of the particular baking soda used was about 6 g of $NaHCO_3$, and 1 level tablespoon of liquid was approximately 10 mL. Students were provided with a reference table of household measure equivalents, as presented in Table 1. The use of non-standard measurement tools was intentional, allowing students to consider measurement uncertainty and experimental limitations.
- Dropper or syringe: a plastic medicine dropper or syringe without needle, marked in mL, to deliver the titrant dropwise. Most students obtained a 5 mL syringe, which allowed reasonably accurate measurement of small volumes and easy counting of drops. (Approximately 20 drops correspond to 1 mL of solution.)
- Small clear container: a transparent glass of approximately 100 mL volume to serve as the titration vessel. This size and transparency

allowed students to observe color changes and gas bubbles, and the glass could withstand the temperature of the hot water bath.

- Personal protective equipment: optional safety gear. Students were encouraged to wear safety glasses, especially when handling hot liquids or pouring vinegar. Gloves and lab coats were not required since all chemicals were food-grade and generally safe, but students could use kitchen gloves or an apron as they felt necessary.

Table 1. Household Measurement Units and Their Approximate Quantities Provided to Students for the Experiment

Household Measure	Description	Approximate Quantity
1 heaped teaspoon – solid	Baking soda powder ($NaHCO_3$)	~6 g
1 tablespoon – liquid	Vinegar or water volume	~10 mL
1 teaspoon – liquid	Solution volume	~5 mL
1 glass– liquid (full)	Small glass/volume	~100 mL
1 mL of solution (for drop counting)	Using a syringe/dropper	~20 drops

Procedure

At the beginning of the activity, students were briefed on safety protocols before the at-home experiment. The primary physical risk involved the use of hot water (approximately 80 °C), which posed a potential burn hazard. Students were instructed to handle heated materials using protective barriers (e.g., cloth or gloves), to work on stable, heat-resistant surfaces, and avoid placing hot liquids near the edge of the workspace. Eye protection was recommended, especially when adding baking soda to hot vinegar, due to the possibility of effervescence and minor splattering. Although no safety incidents occurred, the instructor emphasized minimizing risk. Both vinegar and baking soda are classified as low-hazard household substances; using gloves and lab coats was optional. Adequate ventilation was ensured to prevent CO_2 buildup and to support reaction completion. Students recording the activity for reports were reminded to follow safety guidelines and obtain consent if others appeared in the video. Informed consent was obtained from all participants and confirmed the absence of any allergies to vinegar or cabbage extract. The experiment was conducted safely by all 19 students, with no adverse events reported.

As a preparation task, each student prepared a red cabbage indicator solution and tested it with several household substances to observe the resulting color changes. This activity provided students with initial familiarity with the indicator's behavior, such as its pink color in acidic solutions and green or blue tones in basic ones, before beginning the titration.

At the end of the activity, students were reminded that the reaction mixture consisted primarily of water, sodium acetate (a food-safe salt), and cabbage pigment. Since all components were non-hazardous, the mixtures were safely disposed of by pouring them down the sink with plenty of water. Any leftover baking soda or vinegar could be retained for household use. Equipment (e.g., glasses and syringes) was washed and returned to regular kitchen use.

Throughout the activity, students followed the sequence of steps outlined in the POGIL framework presented below.

Step 1: Pre-Experiment Questions. One week before the activity, students were assigned structured pre-laboratory inquiry questions (as an at-home "pre-lab" homework) to ensure they were ready for the concepts involved. These questions included:

1. The acetic acid concentration in a sample needs to be determined by the titration method. Can sodium bicarbonate be used as a titrant? Please explain your answer.
2. Write the reaction equation between acetic acid and sodium bicarbonate. Do the reaction product(s) have a negative effect on the titration? Please explain your answer. If so, what can be done to minimize the effects?
3. How to calculate a theoretical titration curve of the acetic acid and sodium bicarbonate? Draw a rough sketch of the curve.
4. Can an acid-base indicator be prepared from household materials? Please give examples.
5. How does the color change of the acid-base indicator occur during titration? Please represent it on microscopic, macroscopic, and symbolic levels.

Students submitted brief written answers to these questions via the LMS prior to the lab session. The instructor reviewed the responses and provided collective feedback to the students. The feedback emphasized several key points. First, acetic acid is a weak acid, so using a weak base as the titrant presents challenges unless the reaction is driven to completion. Second, CO_2 formed from H_2CO_3 can affect the pH and must be removed to obtain a clear endpoint. Finally, natural acid–base indicators—such as red cabbage extract—can be effectively used, which many students correctly anticipated in response to question 4.

Step 2: Pre-Observation Prediction Questions and Observations. This step within the POGIL structure represented the implementation phase of the activity. During the implementation of the activity, students joined the live video meeting from their kitchens, with their materials ready. The instructor first discussed safety precautions (emphasizing care with hot water and proper disposal of mixtures) and demonstrated the basic titration setup via camera. Students then followed the procedure stages, which were also outlined on their worksheet. For example, students were initially expected to prepare the necessary solutions, as outlined below. Each student prepared a diluted analyte solution by combining approximately 10 mL of white vinegar with hot water to a total volume of approximately 100 mL. This yielded roughly a 10× diluted vinegar solution at elevated temperature. They also prepared the titrant solution in a second glass by dissolving one heaped teaspoon of baking soda (approximately 6 g $NaHCO_3$) in a full glass of hot water (approximately 100 mL), yielding an approximate sodium bicarbonate solution of 0.7 M (based on assumed mass and volume estimates). Hot water was used in both solutions to facilitate the dissolution of baking soda and to provide thermal energy that promotes CO_2 outgassing during titration. Following the preparation of the required solutions, the first prediction question was: What color change would you expect when adding red cabbage juice to the analyte solution? Based on their preparatory work, most predicted that white vinegar, being acidic (undiluted approximately pH 2–3, diluted approximately pH 3–4), would turn the cabbage indicator a pinkish-red color. Indeed, upon adding the indicator, students observed the solution turn a magenta/pink color, confirming its acidity. Then, students added 3–4 drops of red cabbage indicator solution to the diluted vinegar solution using a syringe and recorded the initial observations. The worksheet then prompted: "What color change do you expect as the titrant is added? What color do you estimate at the endpoint, and what do you think the pH will be at the endpoint?" This encouraged students to consider the titration's progression. Students discussed in groups, and many anticipated that as baking soda solution is added, the solution's color would move from pink towards purple and then to blue if it reaches neutral to slightly basic conditions. Students estimated the endpoint color to be blue. They also reasoned that at the equivalence point for acetic acid with sodium acetate in solution, the pH would be slightly above 7. These predictions were recorded as hypotheses. Students began titrating by adding the sodium bicarbonate solution to the vinegar solution gradually. The procedure followed a two-phase addition approach to save time: first, each student added an initial larger portion of titrant (about one teaspoon, approximately 5 mL) all at once to the acid solution and gently swirled the glass. This was intended to neutralize the bulk of the acid quickly and get the solution closer to the endpoint range. They were cautioned not to add more than this at once to avoid overshooting the endpoint. After this initial addition, they proceeded to add the titrant dropwise using the syringe, while continuously swirling, counting every drop, and observing the solution's color change carefully. They were reminded that approximately 20 drops correspond to 1 mL, to help estimate volume from drop counts. Students were instructed to terminate titrant addition at the first persistent color change indicating the endpoint, defined as a stable transition from pink toward blue after gentle stirring, accompanied by a noticeable reduction in gas bubbling. Students recorded qualitative observations after each increment of about 20 drops of titrant to track the titration's progress, and noted the total volume of titrant added when the endpoint was reached. Representative student observations included: "An initial addition of approximately 5 mL of titrant resulted in a lighter pink solution. Subsequent additions in 20-drop (approximately 1 mL) increments led to a gradual color transition from pink to purple, with continued gas evolution. At around 7 mL total volume, the solution became darker purple and bubbling decreased. Near 7.5 mL, a faint blue color appeared and remained stable upon gentle swirling, accompanied by minimal gas evolution." If any student added one or two drops beyond the first blue appearance, they noted the color became a more pronounced blue, and fizzing stopped entirely, indicating excess base. Students were guided to identify the endpoint at the first persistent blue color that did not revert to pink upon stirring, as this provided the most accurate results. In cases of overshoot, this was addressed during analysis by considering the addition of a small amount of vinegar to test whether the color would revert, analogous to a simple back-titration approach.

Step 3: Thinking About Data Questions. Once the titration endpoint was recorded (each student ended with a solution that was purplish-blue or blue and no more CO_2 evolution), students noted the total volume of titrant used. Subsequently, students engaged in guided data analysis tasks and reflection by addressing the following guided questions:

1. Determine the concentrations of the analyte and titrant solutions used in the titration.

2. Determine the percent by mass of acetic acid in white vinegar you buy from the store (use the given scales and approximate quantities to make the calculations. Assume the density of the vinegar sample to be 1.0 g/mL.)
   To answer this question, which was also the research question of the activity, they were instructed to use the approximate measurements given (from Table 1) for consistency.
3. Compare the percentage of acetic acid written on the label of the white vinegar bottle with the percentage you calculated. Determine the absolute and percent relative errors.
4. Is the percent value of the acetic acid you calculated reasonable? What could be the reasons for the error?
5. How would you calculate the theoretical titration curve of acetic acid and baking soda titration? Use the data table previously transferred to Excel and fill it for the titration curve by pooling data together. Identify the equivalence point and the half-equivalence point from the titration curve. Determine the value of Ka for acetic acid by using the titration curve. If you were asked to identify the appropriate indicator using this titration curve, which indicator would you choose? Use the indicator color and pH range table uploaded to the system to choose an indicator.

The students were expected to fill in the Excel spreadsheet by pooling data together for the fifth question.

Step 4: Post-Experiment Questions. Post-experiment reflective questions were assigned to consolidate conceptual understanding as a post-activity homework assignment and were subsequently discussed in the following week's session.

1. If you used another weak acid instead of sodium bicarbonate, would you get the same titration curve?
2. Suggest an experimental method to determine the percent by mass of lactic acid in a yogurt sample.
3. Match the titration curves with the appropriate acidity constants (titration curves and acidity constants are provided in the worksheet).

This session also served as a platform for discussing the questions posed in the activity worksheet and addressing any additional student inquiries. This comprehensive approach ensured that the learning process was thoroughly completed, allowing for clarification of concepts and deeper engagement with the material. Table 2 illustrates a representative sample of the questions, categorized by instructional stage, and includes expected reasoning to demonstrate how the activity assessed and supported the development of core scientific competencies. These include not only content knowledge but also metacognitive and communicative skills essential for scientific inquiry.

Table 2. Examples of Guided Inquiry Questions from the At-Home Titration Worksheet with Expected Responses

Activity Stage	Sample Guided-Inquiry Question	Expected Reasoning / Response
Pre-Laboratory Preparation	*Write the balanced chemical equation for the reaction between acetic acid and sodium bicarbonate. Do any of the products interfere with titration accuracy? If so, how can this be mitigated?*	CO_2 is a gaseous product that can affect pH. Allowing it to escape (e.g., by heating or not sealing the container) pushes the reaction to completion by shifting the equilibrium (Le Châtelier's principle).
Pre-Experiment Prediction	*How do you predict the color of the solution will change as the titrant is added? What color and pH do you expect at the equivalence point?*	Gradual shift from pink (acidic) toward purple or blue; expected endpoint around pH 7–8, depending on the indicator.
Pre-Experiment Prediction	*What assumptions are we making when using sodium bicarbonate as the titrant, and how might these affect the outcome compared to using a strong base like NaOH?*	Assumes bicarbonate fully reacts and neutralizes acetic acid; however, being a weak base, it may not sharply define the equivalence point, leading to a broader pH transition zone and less distinct endpoint.
During the Experiment	*What do you observe during the addition of the titrant? Identify any temperature change, color shift, or gas evolution. How do these observations support or contradict your predictions?*	Observations include fizzing due to CO_2 release, gradual color change, and slight temperature change. These support predicted chemical changes and indicate progress toward neutralization.
During the Experiment	*How many milliliters of titrant were required to reach the endpoint? How consistent are your results with those of your peers? What might explain any discrepancies?*	Variations may be due to drop size inconsistency, endpoint misjudgment, or differences in vinegar concentration. Encourages discussion of procedural error and technique variation.

Table 2. (Continued)

Activity Stage	Sample Guided-Inquiry Question	Expected Reasoning / Response
Data Analysis	*Calculate the percent by mass of acetic acid in your vinegar sample. Compare it to the label value and determine both absolute and percent error. Are the differences scientifically significant?*	Expected concentration approximately 4–5%; acetic acid (w/v) errors within ±10–15% are acceptable. Deviations beyond this may indicate measurement issues, endpoint detection errors, or impurities.
Data Analysis	*Was your titration curve linear, exponential, or sigmoidal in nature? Use your data to sketch the shape and interpret the location of the inflection point.*	Curve is sigmoidal, with a sharp rise at the equivalence point. Inflection point corresponds to rapid pH change and indicates the stoichiometric endpoint.
Error Evaluation	*Identify possible sources of systematic and random error in your procedure. How could these be reduced in future repetitions of the experiment?*	Systematic: miscalibrated tools, wrong concentration assumption. Random: inconsistent drop size, subjective color interpretation. Improvements: use standard volumes, better lighting, and replicate trials.
Post-Experiment Extension	*If you used a different indicator (e.g., phenolphthalein instead of cabbage extract), how might the visual endpoint change? What does this imply about the indicator's pKa and color range?*	Phenolphthalein changes near pH 8–10, so the endpoint might appear later and shift the perceived equivalence point. Cabbage extract has a broader pH range but less precision. Indicator choice affects color transition and data interpretation.

Data Collection Instruments

Data for this study were collected from 19 student participants using multiple sources: (1) POGIL activity worksheets and group discussion transcripts supplemented by instructor observation notes. Each student completed a structured POGIL worksheet that included procedural questions, conceptual questions, guided prediction prompts, spaces for observation notes, and post-lab calculation and reflection questions. These worksheets captured individual student responses during and after the activity. Furthermore, qualitative observational data were gathered by the instructor during the live sessions – for example, notes on group discussions in breakout rooms and the points that student spokespersons shared during whole-class debriefings. To ensure confidentiality, all student-written responses were anonymized by removing any identifying information before analysis.

Data Analysis

Both quantitative and qualitative analysis techniques were employed. For the quantitative data, each student's titration calculation was reviewed: from the volume of $NaHCO_3$ solution they used at endpoint, the moles of $NaHCO_3$ and thus the moles of acetic acid neutralized were determined, and the concentration (and % w/v) of acetic acid in the original vinegar was calculated. These values were compiled, and basic descriptive statistics such as range, mean, standard deviation were computed to characterize the overall distribution of results. This allowed evaluation of the accuracy and consistency of the at-home titrations.

For the qualitative data, an inductive thematic analysis was conducted on students' written explanations and the transcribed portions of group discussions. The goal was to identify common themes in students' reasoning, misconceptions, and conceptual breakthroughs. Initially, two broad categories of outcomes were considered: (1) Conceptual understanding related to acid–base chemistry and equilibrium, and (2) Process or inquiry skills development. Within these categories, specific codes were generated based on the data. For example, under conceptual understanding, codes included "recognition of CO_2 removal shifting equilibrium," "understanding of weak vs strong titration differences," "indicator color interpretation," and "connecting to real-world examples." Under process skills, codes included "team collaboration evidence," "experimental uncertainty awareness," and "student engagement level." Two researchers independently coded a subset of the data to ensure reliability, then discussed and refined the coding scheme. The coded qualitative data were then examined for patterns and representative examples. Triangulation was achieved by comparing insights from the worksheets and discussion transcripts – for instance, if a concept was mentioned in group discussion and also showed up in many students' written answers, it was considered strongly evidenced.

Ethical Considerations

This study was conducted after all participants provided informed consent. Given the at-home nature of the experiment, additional precautions were taken to ensure safety and privacy. Students confirmed that they had no known allergies to any materials as a formality before beginning. All participants were briefed on safe use of hot water, and agreed to follow the safety guidelines provided.

Results and Discussion

Titration Outcome and Accuracy

All 19 participants successfully completed the titration procedure and determined the acetic acid content of their white vinegar sample. The measured acetic acid concentrations in vinegar samples ranged from approximately 2.6% to 5.5% (w/v). The mean measured concentration was 3.8% w/v with a standard deviation of 0.83 percentage points. Given the simple tools and potential sources of error in the home environment (e.g., imprecise volume measurement using kitchen utensils, human color perception in judging the endpoint), this level of accuracy can be considered acceptable within the constraints of non-standardized, home-based experimental conditions. Notably, the class mean (3.8% (w/v)) falls within the typical label range for white vinegar (4–5% acetic acid (w/v)), and no individual result deviated by more than approximately 1.3 percentage points.

Outliers were observed at both extremes of the distribution: a small subset of students obtained low values (approximately 2.6–2.8% (w/v)), and a small subset of others obtained high values (approximately 5.2–5.5% (w/v)). These deviations provided a productive basis for analytical discussion. The low values were likely due to undershooting the endpoint. In these cases, some acetic acid remained unneutralized, leading to an underestimation of its concentration. On the other hand, the highest results were likely due to overshooting the endpoint. Students were able to relate these errors to their technique: during the debrief, those who obtained the lowest values reported that they were hesitant and may have stopped adding titrant at a persistent purple rather than blue, whereas those with the highest values indicated that they might have added a few too many drops beyond the first sign of blue.

This discussion reinforced an important experimental lesson about titration technique: the need for careful, dropwise control near the endpoint and the value of a subtle color change as the stopping point. We also discussed a potential refinement: if a student suspected they overshot, they could add a drop of vinegar to see if the solution color shifts back (from blue toward purple), which is analogous to a back-titration to confirm the endpoint. Such techniques are commonly used by chemists to improve accuracy, and introducing this idea helped students think critically about how to refine experimental procedures.

It is noteworthy that while the color change to blue was used as the primary indicator of neutrality, many students also observed the cessation of CO_2 fizzing as a complementary indicator of the endpoint. As one student described in her activity sheet, “When the solution turned blue, I noticed the bubbles stopped almost completely – it was like the reaction went quiet.” This observation aligns with what Nguyen et al. (2023) reported in their at-home citric acid titration: the end of gas evolution can reliably signal the endpoint in a weak acid–weak base titration. In this study, since the cabbage indicator provided a clear visual cue, we did not rely solely on gas cessation; however, the concurrence of these two signs (blue color and no bubbles) strongly reinforced students’ understanding that chemistry often provides multiple indicators of a reaction’s progress. This provided clear experiential evidence that neutralization was essentially complete complete once CO_2 evolution ceased.

Overall, the quantitative titration findings indicate that the home-based experiment yielded results with moderate accuracy and limited precision. Despite observable deviations from the expected stoichiometric ratios, the mean percent error for calculated acetic acid concentration remained within an acceptable pedagogical margin (e.g., 10–15%), consistent with previous home-based titration studies (Andrews et al., 2020). The standard deviation across replicates was notably higher than that observed in traditional laboratory settings, reflecting procedural inconsistencies such as imprecise volume measurements, incomplete neutralization, or variations in household measuring tools.

Rather than diminishing the value of the experiment, this variability functioned as a productive learning opportunity. Unlike prescriptive “cookbook-style” laboratory exercises that often yield uniform results, this inquiry-based activity produced a range of outcomes, prompting students to engage in interpretive reasoning. The variability introduced them to the nature of experimental uncertainty and led to reflective classroom discussions focused on precision, accuracy, and error classification. Students critically examined potential sources of deviation, including differences in measurement techniques, subjective interpretation of indicator color changes, and overshooting or undershooting the endpoint. Peer comparisons encouraged exploration of systematic versus random errors, fostering a more nuanced understanding of scientific measurement. As one student group noted, “Not everyone got exactly 5.0% (w/v), so we had to consider what might have caused the differences,” initiating an authentic inquiry into the reliability of data. In this way, the home-based titration experiment not only supported the development of conceptual understanding in acid–base chemistry but also cultivated key scientific process skills. It offered a realistic representation of the inherent variability in empirical data collection and emphasized the importance of evidence-based reasoning, an essential component of scientific literacy.

Conceptual Understanding and Equilibrium Implications

Beyond the numerical results, the activity provided rich qualitative insights into student understanding of acid–base chemistry and equilibrium. One of the most important conceptual insights for students was understanding why the titration was feasible despite the use of a weak base as the titrant. Initially, many students held the standard idea (taught in lecture) that a weak acid should not be titrated with a weak base because the reaction would not go to completion and the endpoint would be ill-defined. This was evident in their pre-lab answers: when asked if $NaHCO_3$ could be used to titrate acetic acid, most answered “no,” citing the lack of a sharp pH change and incomplete neutralization, which is indeed the usual rationale in analytical chemistry courses. This reflected a sound general understanding but also a conditional misconception – they assumed it was impossible to get a clear endpoint, not yet considering the role of removing CO_2. This created cognitive tension, as students were about to carry out an activity they had initially considered impossible. A few students speculated in advance that “maybe if something like CO_2 were removed it could work,” revealing an intuitive understanding observed in only a small number of students, suggesting a more advanced prior understanding of equilibrium.

Through the guided inquiry process and the experiment itself, students deduced the critical role of carbonic acid and CO_2 in this system. In the post-lab reflections, most students articulated that the formation of CO_2 gas (from

H_2CO_3) and its escape was driving the reaction to completion. Essentially, students demonstrated an understanding of Le Châtelier's principle in this context. As one student explained, "Because CO_2 gas leaves the solution, the reaction keeps shifting to make more products, so eventually almost all the acetic acid reacts." This explains why the weak base (bicarbonate) can effectively neutralize the weak acid under these conditions. This insight demonstrated a deepened conceptual understanding of chemical equilibrium – they connected a macroscopic observation (bubbling gas) to a microscopic equilibrium shift. Importantly, this directly addresses the initial misconception: students learned that a weak base can be used to titrate a weak acid if a product (CO_2) is continuously removed. The activity thus turned an "impossible titration" scenario into an inquiry in which students identified the condition that made it possible. This provides evidence of conceptual change facilitated by the experiment: students had to reconcile their prior knowledge that a weak acid and a weak base do not react to completion with new evidence showing that the reaction did go to completion because CO_2 was removed. This outcome aligns with findings by Salame et al. (2022), who reported that students often struggle to connect macroscopic observations with equilibrium concepts in acid–base reactions. By directing attention to equilibrium shifts, the activity supported students in bridging this conceptual gap.

Another conceptual area strengthened was understanding pH indicators and endpoint detection. Students experienced firsthand that a natural indicator like red cabbage has a broader transition range compared to a synthetic indicator like phenolphthalein. During the titration, they observed the solution transition through a gradient of colors (from pink to purple to blue) rather than a single sudden change. Initially, about 60% of the students predicted the endpoint color correctly (blue), while around 30–40% thought it might end up purple (assuming neutral pH 7 would be the stopping point). After the experiment, they recognized that the actual endpoint was slightly basic, hence the cabbage indicator turned greenish-blue. It was noted that if phenolphthalein had been used, the endpoint would have been a faint pink appearing at approximately pH 8, which follows the same basic principle. This helped solidify their understanding that endpoint is not necessarily pH 7 for acid–base titrations involving weak acids – a common misconception is that "neutralization = pH 7," which they observed does not necessarily hold for systems involving weak acids (the equivalence point pH is >7 if the base is strong, or here slightly >7 because the conjugate base present is weakly basic). By observing and talking through the color changes, students gained an appreciation for indicator selection: one student asked, "Would phenolphthalein have worked here?" – leading to a productive discussion on indicator pH ranges. (Answer: Phenolphthalein, which changes around pH 8–10, likely would have worked, since our endpoint was approximately pH 8; it would have given a pink endpoint. Methyl orange, by contrast, would not have been suitable since its range is around pH 3–4, too low for this titration.)

These observations were also connected to a titration curve. During the post-lab session, the class sketched a simplified titration curve for acetic acid with $NaHCO_3$. Using pooled data, it was noted the starting pH (approximately 3), a gradual rise (buffer region), a less steep equivalence point around approximately pH 7 (if CO_2 is removed, otherwise lower), and a final plateau in the basic range if excess base is added. Students realized that by removing CO_2, the curve was effectively shifted upwards in the final portion, allowing it to pass through the neutral range into basic – otherwise, CO_2 accumulation would have buffered the solution around pH 6, and the curve might plateau before reaching 7. This exercise reinforced their conceptual grasp by visualizing the titration progress in terms of pH vs. volume and relating it back to what they observed. In essence, the experiment served as a concrete anchor for abstract concepts like titration curves and equilibrium shifts, exemplifying Kolb's experiential learning cycle (Kolb et al., 2014) from concrete experience to abstract conceptualization.

Finally, students were able to connect the experiment to real-world contexts, demonstrating an enriching of their conceptual framework. For example, when discussing why we heated the solution to help CO_2 escape, one student spontaneously mentioned, "This is like how soda goes flat faster when warm." This observation connected Henry's law and gas solubility to a familiar experience – recognizing that warm soda loses CO_2 quickly just as our hot solution lost CO_2 to drive the reaction. Another connection was made to environmental chemistry with the question, "What do you think the pH of rainwater is, and why?" Students recalled that rainwater is slightly acidic (around pH 5.6) because CO_2 in the atmosphere dissolves into raindrops, forming carbonic acid. Immediately, they observed the parallel: dissolved CO_2 made our titration solution acidic until it was driven out; similarly, atmospheric CO_2 makes natural rain slightly acidic. These observations indicate that students were integrating conceptual understanding and recognizing its relevance beyond the laboratory context. This form of knowledge transfer is widely regarded as an indicator of meaningful learning (Taber, 2014), as students moved beyond procedural execution to internalizing and applying concepts to everyday phenomena.

Analysis of Student Responses

The guided-inquiry format prompted students to articulate their reasoning at each stage, thereby revealing both misconceptions and learning gains. For instance, during the feedback phase on students' responses to the preparatory questions, many provided incomplete definitions of what it means for acetic acid to be a "weak" acid, such as "it doesn't ionize much" or "it doesn't dissociate." This oversimplification was addressed in class, clarifying that a weak acid does ionize but only partially, establishing an equilibrium. Likewise, initially, most students argued that sodium bicarbonate could not be used to titrate acetic acid, because they remembered that weak acid–weak base titrations are not performed in standard practice. They mentioned the lack of a sharp pH change and incomplete reaction. However, after the experiment, their written conclusions unanimously noted that it was possible under the condition of CO_2 removal. This shift from pre- to post- activity responses indicates the instructional impact. It highlights how the activity fostered a form of guided rediscovery, in which students began with conventional assumptions, encountered results that challenged those assumptions, and subsequently constructed their own explanations with the support of instructional scaffolding. In educational terms, this is a potent way to address alternative conceptions (Salame et al., 2022). Our observations align with Salame and Montero's findings that students often rely on rote rules (e.g., "never titrate weak with weak"); by challenging that rule through inquiry, students were forced to think conceptually about why that rule exists and how to overcome its limitations.

All students were able to write the balanced overall reaction: $CH_3COOH(aq) + NaHCO_3(aq) \rightarrow CH_3COONa(aq) + H_2CO_3(aq)$. And nearly all knew H_2CO_3 decomposed to $CO_2(g) + H_2O(l)$. When discussing whether the products have a disturbing effect on titration, about 80% identified CO_2 as an interfering factor that would affect pH and thus the endpoint. Solutions they proposed for mitigation included "remove the CO_2 by heat or letting gas out," and one student even said, "maybe use a closed container with a balloon to capture CO_2". These answers show that students were thinking critically about the reaction's equilibrium.

All students realized red cabbage could be one (some had used it in high school before). Other students mentioned turmeric (turns red in base), blueberry juice, hibiscus tea, and even flower petals. Indeed, there are many natural pH indicators. Since red cabbage had already been assigned as the indicator, students' alternative suggestions were acknowledged, and it was noted that these could be explored in future activities.

All students correctly answered the question regarding the initial color of the indicator, identifying the vinegar solution with the indicator as pink/red, consistent with its acidic nature. Their responses reflected an understanding of the behavior of cabbage-derived pigments in acidic conditions.

The question predicting endpoint color was a bit trickier: approximately 60% predicted blue (correct for slightly basic endpoint). Approximately 30% said purple (thinking neutral would be purple). This led to a discussion: the exact color of red cabbage at pH 7 can be a bit ambiguous – purple or blue – but given the basic salt (acetate) likely pushes the endpoint >7, blue was observed. Students who identified the solution as purple were not entirely incorrect from a conceptual standpoint, as a neutral solution can indeed present a purple hue with the indicator used. However, in the practical execution of the experiment, the observed endpoint was mildly basic, which corresponded to a blue color. This discrepancy highlights the importance of precise experimental techniques and careful observation in determining the actual reaction endpoint. This was explained by noting that the choice of indicator is important: an indicator's transition range matters. Red cabbage has a broad range; our approach relied on noticing a distinct shift rather than a single sharp change. This was an opportunity to talk about indicator ranges and the concept of a titration curve. To support conceptual understanding, a simplified virtual titration curve of acetic acid titrated with $NaHCO_3$ was presented. The curve began around pH 3, increased gradually, rose more steeply near the equivalence point (estimated at approximately pH 8 due to acetate formation), and eventually plateaued in the basic range upon addition of excess base. It's not as steep as a strong/strong titration, but because CO_2 escapes, it does rise. If CO_2 stayed, the curve would buffer around pH 6 and never get to 7+. Students found it notable that our method basically altered the titration curve in situ by removing CO_2.

An important conceptual question in the worksheet asked what gas is being released and why its release helps drive the neutralization reaction forward. Students answered: "The gas being released is CO_2, and solubility decreases with heat, so using hot water kept the solution hot and let CO_2 leave. According to Le Châtelier's principle, the decrease in gas shifts the equilibrium to the right, which is correct. Some also mentioned that physically, the bubbles carried the CO_2 out. It was emphasized that, even in the absence of deliberate heating, the reaction generates heat through exothermic neutralization and gas expansion and produces agitation due to bubbling;

however, external heating significantly enhances the effectiveness of CO_2 removal.

Many noted that with NaOH, there would be no gas, and phenolphthalein could be used to see a sharp pink endpoint, typically with much less volume needed (since NaOH is usually approximately 0.1 M in school labs, not 0.7 M as the baking soda, but one would not dilute vinegar as much either).

Some students recognized that the equivalence point pH would differ depending on the titrant used. For instance, a titration involving sodium hydroxide (NaOH) would typically yield an equivalence point pH greater than 8, as the acetate ion (CH_3COO^-), the conjugate base of acetic acid, is a weak base that hydrolyzes in water to produce hydroxide ions. In contrast, the titration employing sodium bicarbonate ($NaHCO_3$) as the titrant resulted in an equivalence point pH in the range of 7-8. This outcome is attributed to the specific experimental setup, where the effective removal of carbon dioxide (CO_2) during the titration caused the bicarbonate ion (HCO_3^-) to behave akin to a hydroxide ion (OH^-). In this context, the HCO_3^- effectively promoted the reaction toward completion under conditions where CO_2 was continuously removed, leading to a near-neutral to slightly basic equivalence point. This observation highlights how experimental conditions can significantly influence the observed chemical behavior and the resulting pH at the equivalence point. This represents an advanced observation, linking the role of HCO_3^- as the conjugate base of a weak acid (H_2CO_3) to system conditions. In a closed system, the reaction would not proceed fully; however, under open conditions, continuous CO_2 removal shifts the equilibrium, allowing the reaction to proceed more effectively.

Addressing Alternative Conceptions

The activity allowed us to uncover and address some misconceptions:

Equating “Weak” with “Ineffective”: Initially, students thought a weak base could never neutralize all of a weak acid. It was shown that by altering conditions, it can. This nuances their understanding of “weak” – it’s about the degree of dissociation/equilibrium, not an absolute inability to react. Many were surprised that in the end, essentially all the acetic acid was neutralized; one said, “I thought there would always be some acetic acid left unreacted, but we practically got rid of it by boiling off the CO_2 – that’s clever.”

Indicator Function: A few students were unsure if the indicator color change was reversible or if it “uses up” the indicator. One asked why the color didn't change back when the solution cooled. This opened a discussion about the indicator color change being due to pH, not temperature (though extreme temperature can sometimes shift the equilibrium of the indicator slightly). So the reason it stayed blue is because the solution's composition (acetate ions present, no free acid) maintained a basic pH, not because heating or cooling directly affected the pigment's color.

Calculation Pitfalls: Some students initially forgot to account for the dilution of vinegar or made a misstep in unit conversion (drops to mL). These were resolved in group work. One group had a low outlier result because they mistakenly used 100 mL as the vinegar volume in the calculation instead of 10 mL. When the class plotted all reported % (w/v) acetic acid, that group's result (approximately 0.5%) was obviously wrong, prompting them to double-check and find the error.

Critique of Student Answers

In students' notes, a student wrote, "initial solution pH maybe around 5 – purple color." This is slightly misconceived: vinegar diluted 1:10 likely has a pH closer to 3-4, which normally yields a pink, not purple, in cabbage indicator. It's possible the indicator was a bit dilute or the student's color distinction was off. This indicates a point for improvement: To enhance students' understanding of the cabbage indicator's behavior, a color reference chart for the indicator could be provided in the materials. Indeed, one or two groups did a quick "indicator test," and it helped them recognize the colors.

Another interesting student answer came in response to: "Why did we dilute the vinegar sample (adding 10 mL of vinegar to approximately 90 mL water) before titrating?" This was not explicitly asked in the worksheet, but was raised by a curious student. The instructor addressed it verbally: we diluted to use a larger volume so the color change would be easier to see in the larger volume, and also because using full-strength vinegar (0.83 M) with 0.714 M $NaHCO_3$ might require almost equal volumes; dilution spreads it out and gives a more gradual change, plus avoids too vigorous a reaction all at once. In retrospect, adding this as a question in the worksheet could be good, as it could prompt students to consider the rationale for dilution (which is often done in titrations to get better precision or to accommodate a convenient titrant volume). Most accepted the explanation once given. It touches on experimental design: choosing concentrations so that titrant volume will be reasonable (approximately 10 mL in our case, rather than 1 mL or 100 mL extremes).

In the assessment of student laboratory reports, approximately half of the class demonstrated a clear comprehension of equilibrium shifting due to the removal of a gaseous product. This indicates a successful transmission of this key concept. The following students' excerpts highlight this understanding:

Heating the solution was essential to reduce CO_2 solubility (Henry's law) and thus push the reaction forward. This made our weak base titration behave like one with a strong base."

"The weak acid-weak base titration succeeded because one product (CO_2) left. If it hadn't, the endpoint would be very hard to detect and incomplete."

"This experiment taught me why titrations usually avoid weak base titrants – but also how to overcome that issue creatively."

These reflections show a good grasp of the core chemical concept, and they appropriately credit the method.

On the other hand, a few students in their conclusions still referred to $NaHCO_3$ as if it directly neutralized all the acid by being a base, without mentioning the role of CO_2 removal. This suggests that not everyone fully internalized the equilibrium aspect; they might just think, "We added enough baking soda, so of course, it neutralized." To encourage deeper engagement and critical thinking regarding the observed phenomenon, the following question could be posed for further discussion or reflection: "If you tightly covered the reaction cup so no gas escapes, what do you predict would happen? Would the color ever turn blue?" This hypothetical could force them to articulate the necessity of CO_2 escape. Additionally, including the reference to the work of Nguyen et al. (2023) in class discussion helped validate for students that this approach is documented – they were intrigued to learn that others have titrated citric acid with baking soda and even compared it to a traditional titration. It made them feel part of a broader scientific conversation.

Inquiry and Pedagogy Reflections

From an instructional perspective, the open-ended, inquiry nature of the activity proved effective. Students were initially challenged by being asked a research-oriented question regarding the acetic acid percentage in vinegar, which they initially assumed might have a predefined answer. A few admitted later, "I was tempted to just Google the vinegar acidity." If the question were phrased plainly, it could encourage reliance on external sources by looking up typical vinegar strength. To address this and promote authentic inquiry, the task was intentionally framed to require experimental determination and comparison with labeled values. In future implementations, an alternative formulation could further strengthen inquiry by asking whether the manufacturer's stated acidity aligns with experimentally determined values. This way, the question is more open and cannot be answered without doing the experiment and considering experimental error. Alternative formulations may also prompt consideration of error sources and variability, thereby avoiding reliance on a single expected value. By performing the titration and associated calculations, students took ownership of their result. The collaborative component also played an important role. Although each student conducted an individual titration, results were compared within groups. Variations in results (e.g., 4.5% vs. 5.3% w/v) prompted discussion of potential causes. Students considered factors such as endpoint detection, dropwise addition accuracy, and the potential benefits of replicate trials and averaging. This reflects the "open-ended results" aspect: not everyone got an identical number, so they had to interpret a range of results. This reflects authentic scientific practice, where variability in data is expected. This prompted discussion of precision and accuracy. Students were also encouraged to propose strategies for improving accuracy, including conducting replicate titrations and averaging results, using digital tools (e.g., a smartphone-based pH measurement) for more quantitative endpoint detection and practicing with standard solutions. These responses indicate a developing understanding of experimental design.

Inquiry-Based Learning Benefits

Students generally appreciated the POGIL format, despite its novelty. This approach required active engagement through guiding questions, encouraging students to deduce the rationale behind the experimental procedure rather than passively follow predefined instructions. This shift supported a deeper understanding of the scientific process and underlying laboratory principles. For example, when one student asked whether all the baking soda could be added at once, a peer explained that gradual addition was necessary to accurately identify the endpoint and to control CO_2 evolution. This peer-to-peer explanation occurred without instructor intervention, demonstrating that the inquiry prompts supported independent reasoning and collaborative sensemaking.

Connection to Student Learning Outcomes

Students applied stoichiometry and solution chemistry calculations in an authentic experimental context. Rather than solving textbook problems about acid-base titration, students performed calculations based on their own experimental data. This contextual approach may support improved retention. Students gained familiarity with a natural indicator, recognizing both its limitations (e.g., broad color transition range) and advantages (e.g., accessibility and visual engagement) compared to synthetic indicators. They also engaged with underlying buffer systems, including the acetic acid/acetate and carbonic acid/bicarbonate equilibria. Some students observed that near the endpoint, incremental additions produced only subtle changes, prompting questions about the transition to more rapid color change. These observations suggest emerging understanding of buffering capacity and equilibrium behavior. The activity also supported the development of laboratory skills, including measurement, careful observation, and safety awareness in a non-traditional environment. Students reported that conducting experiments in a home environment increased engagement while also raising awareness of safety considerations outside the traditional laboratory setting.

Adaptation and Extension

The success of this activity suggests several promising directions for future adaptation. Given the accessibility of materials and the relative safety of the procedure, the experiment has potential for implementation in secondary-level chemistry education to support instruction on acid–base reactions and indicator use. Furthermore, the activity can be meaningfully extended to environmental chemistry contexts. As previously mentioned, students could, for example, collect rainwater samples and qualitatively investigate acidity through titration with baking soda. Although the low ionic strength of rainwater limits precise quantification, the concept of carbon dioxide forming carbonic acid in natural waters could be explored conceptually. Similarly, CO_2 evolution observed in this experiment can be compared to ocean acidification processes, where atmospheric CO_2 dissolves in seawater and forms carbonic acid—representing a reverse process in which CO_2 is absorbed rather than released.

Additional adaptations may involve exploring alternative indicators or analytes. For instance, students could investigate the titration of other household acids such as lemon juice, orange juice, or carbonated beverages using baking soda. While the presence of multiple acids or buffering agents in these substances introduces complexity, it may also present opportunities for deeper inquiry. A particularly innovative modification, inspired by Nguyen et al. (2023), involves the use of a nucleation-based approach (e.g., addition of Mentos) to promote rapid CO_2 release. This technique may provide a practical means for enhancing gas removal and improving endpoint detection.

A dedicated discussion of acceptable error in real-world product analysis would also be beneficial. This would bridge the gap between theoretical laboratory exercises and practical applications, highlighting the importance of error analysis in quality control and industrial settings. In an actual analytical lab, even a 5% error might be considered high; however, given the tools and equipment utilized in this activity, higher error margins appear reasonable. This outcome highlights the limitations of home-based experimentation and provides a valuable opportunity to discuss the importance of proper laboratory instruments for industrial applications and other professional settings.

Conclusion

In summary, the at-home POGIL-based vinegar titration activity engaged students in an inquiry-driven experiment using accessible household materials to address a relevant chemical question: the determination of acetic acid content in vinegar. Through the guided inquiry framework, students developed procedural competence and deeper conceptual understanding of acid–base reactions, chemical equilibrium, and indicator behavior. Notably, students recognized how a weak base (sodium bicarbonate) can effectively neutralize a weak acid, illustrating the application of Le Châtelier's principle through CO_2 removal.

The activity also addressed common misconceptions in acid–base chemistry, particularly the assumption that neutralization necessarily results in pH 7. Through direct observation of a weak acid–base titration, students recognized that neutralization does not necessarily yield a neutral pH and that equilibrium can be shifted toward completion through removal of reaction products (e.g., CO_2).

From a pedagogical perspective, this activity highlights the value of integrating guided inquiry with at-home laboratory experiences. Students were not merely following prescriptive procedures but were actively engaged in interpreting phenomena and drawing conclusions based on evidence. The open-ended nature of the outcomes, reflected in variation across student results, created opportunities for discussion of experimental uncertainty, procedural differences, and chemical reasoning. This approach aligns with the goals of modern chemistry education, which prioritize critical thinking and conceptual understanding over confirmation of expected values. Moreover, the collaborative structure inherent in the POGIL framework encouraged students to articulate reasoning and learn from peers. The variation across student results further created opportunities for discussion of experimental uncertainty, procedural differences, and chemical reasoning. Notably, remote implementations of POGIL have also been shown to support social interaction and collaborative learning, both of which offer additional benefits for student learning (Gao et al., 2025; Howley, 2020), reinforcing the potential of guided inquiry models even in nontraditional instructional settings.

Overall, the at-home POGIL titration of vinegar demonstrates how remote learning, when combined with guided inquiry, can support effective chemistry learning. In comparison with existing studies, the integration of structured guided inquiry contributes a distinct pedagogical dimension, demonstrating how the inquiry-based framework enhances both conceptual understanding and student engagement. Students learned not only how to perform a titration using nontraditional materials but also how to interpret experimental outcomes through an inquiry-based framework. These findings support the continued development of student-centered, inquiry-driven at-home laboratory activities grounded in core scientific concepts. The approach is readily adaptable across diverse educational settings, including both secondary and higher education contexts. The aim is to enrich their existing chemistry curricula by incorporating its pedagogical benefits.

Future research may further evaluate and extend this approach. For example, comparative studies could examine differences between guided inquiry and traditional instructional approaches. Additionally, pre- and post-instruction assessments could be used to quantify learning gains (such as an acid–base conceptual test or an attitude survey). Broader implementation across multiple institutions would help assess generalizability of the approach; such studies could reveal how well the POGIL at-home titration works in different contexts. As the participants were pre-service chemistry teachers, an interesting extension is to examine how this experience influences their teaching perspectives – for example, would they be more likely to use inquiry-based experiments in their future classrooms?

While the activity was successful as implemented, several refinements could further enhance its effectiveness in future iterations. For example, providing a color reference chart for the red cabbage indicator may improve endpoint identification. Incorporating digital measurement tools may further enhance quantitative accuracy. For instance, students could use a smartphone app or even simply photograph the solutions and use a color-analysis tool to get an objective measure of the color change. This could provide additional data like RGB values correlating with pH and allow more precise monitoring of the titration's progress, complementing their visual observations and making the experiment more quantitative. Encouraging replicate titrations would reinforce reproducibility and improve confidence in results. Due to class time constraints, most students only did a single trial in this study, but even one repeat could help identify if an outlier result was due to technique error.

Disclaimer

None

References

Andrews, J. L., de Los Rios, J. P., Rayaluru, M., Lee, S., Mai, L., Schusser, A., & Mak, C. H. (2020). Experimenting with at-home general chemistry laboratories during the COVID-19 pandemic. *Journal of Chemical Education,* 97(7), 1887–1894. https://doi.org/10.1021/acs.jchemed.0c00483

Atkins, P., & de Paula, J. (2022). *Atkins' physical chemistry* (12th ed.). Oxford University Press.

Bretz, S. L. (2013). Evidence for the importance of laboratory courses. *Journal of Chemical Education,* 90(3), 293–294. https://doi.org/10.1021/ed300384g

Bruce, M. R., Bruce, A. E., Bernard, S. E., Bergeron, A. N., Ahmad, A. A., Bruce, T. A., Perera, D. C., Pokhrel, S., Saleh, S., Tyrina, A., & Yaparatne, S. (2021). Designing a remote, synchronous, hands-on general chemistry lab course. *Journal of Chemical Education,* 98(10), 3131–3142. https://doi.org/10.1021/acs.jchemed.1c00559

Caraballo, R. M., Medina, L., Gomez, S., Vensaus, P., & Hamer, M. (2021). Turmeric and RGB analysis: A low-cost experiment for teaching acid–base equilibria at home. *Journal of Chemical Education,* 98(3), 958–965. https://doi.org/10.1021/acs.jchemed.0c01165

Chang, R., & Goldsby, K. (2021). *Chemistry (*13th ed.). McGraw-Hill Education.

Farrell, J. J., Moog, R. S., & Spencer, J. N. (1999). A guided inquiry general chemistry course. *Journal of Chemical Education,* 76(4), 570–574. https://doi.org/10.1021/ed076p570

Gao, Q., Tong, M., Sun, J., Zhang, J., Liu, Y., & Zhou, Y. (2025). A study of process-oriented guided inquiry learning (POGIL) in the blended synchronous science classroom. *Journal of Science Education and Technology,* 34, 103–121. https://doi.org/10.1007/s10956-024-10155-3

Harrison, R. M., & De Mora, S. J. (1996). *Introductory chemistry for the environmental sciences* (2nd ed.). Cambridge University Press.

Howley, I. (2020). Adapting guided inquiry learning worksheets for emergency remote learning. *Information and Learning Sciences,* 121(7/8), 549–557. https://doi.org/10.1108/ILS-04-2020-0087

Hunnicutt, S. S., Grushow, A., & Whitnell, R. (2015). Guided-inquiry experiments for physical chemistry: The POGIL-PCL model. *Journal of Chemical Education,* 92(2), 262–268. https://doi.org/10.1021/ed500421d

Kolb, D. A., Boyatzis, R. E., & Mainemelis, C. (2014). Experiential learning theory: Previous research and new directions. In P. R. Sternberg & R. K. Sternberg (Eds.), *Perspectives on thinking, learning, and cognitive styles* (pp. 227–247). Routledge.

Lanni, L. M. (2014). Filling a plastic bag with carbon dioxide: A student-designed guided-inquiry lab for advanced placement and college chemistry courses. *Journal of Chemical Education,* 91(9), 1390–1392. https://doi.org/10.1021/ed500058v

Levie, R. de. (1991). *Aqueous acid–base equilibrium and titrations.* Oxford University Press.

Lewis, S. E., & Lewis, J. E. (2005). Departing from lectures: An evaluation of a peer-led guided inquiry alternative. *Journal of Chemical Education,* 82(1), 135–139. https://doi.org/10.1021/ed082p135

Li, J., O'Neill, M. L., Pattison, C., Zhou, J. H., Ito, J. M., Wong, C. S., Yu, H.Z., & Merbouh, N. (2023). Mobile app to quantify pH strips and monitor titrations: Smartphone-aided chemical education and classroom demonstrations. *Journal of Chemical Education,* 100(9), 3634–3640. https://doi.org/10.1021/acs.jchemed.3c00227

Moog, R. S., & Spencer, J. N. (2008). *Process-oriented guided inquiry learning (POGIL).* Oxford University Press.

Moog, R. S., Creegan, F. J., Hanson, D. M., Spencer, J. N., Straumanis, A., Bunce, D. M., & Wolfskill, T. (2008). POGIL: Process oriented guided inquiry learning. In R. S. Moog & J. N. Spencer (Eds.), Process-oriented guided inquiry learning (ACS Symposium Series, Vol. 994, pp. 1–13). *American Chemical Society.* https://doi.org/10.1021/bk-2008-0994.ch002

Nguyen, J., Akdeniz, A., Anfuso, C. L., & Morris, J. D. (2023). A simple at-home titration: Quantifying citric acid in lemon juice with baking soda and Mentos. *Journal of Chemical Education,* 100(2), 739–744. https://doi.org/10.1021/acs.jchemed.2c00205

Parel, P., Burnett, L., Geoffroy, M., Parel, J., & Hao, L. (2021). Determining the acetic acid concentration in white vinegar: An at-home undergraduate chemistry experiment during the COVID-19 pandemic. *ChemRxiv.* https://doi.org/10.26434/chemrxiv-2021-hxb4r

Petrucci, R. H., Herring, F. G., Madura, J. D., & Bissonette, C. (2007). *General chemistry: Principles & modern applications* (9th ed.). Pearson Prentice Hall.

Ruder, S. M., & Hunnicutt, S. S. (2008). POGIL in chemistry courses at a large urban university: A case study. In R. S. Moog & J. N. Spencer (Eds.), Process-oriented guided inquiry learning (*ACS Symposium Series,* Vol. 994, pp. 75–88). American Chemical Society.

Salame, I. I., & Montero, A. (2022). Examining some of the students' challenges and alternative conceptions in learning about acid–base titrations. *International Journal of Chemistry Education Research,* 6(1), 1–9. https://doi.org/10.20885/ijcer.vol6.iss1.art1

Taber, K. S. (2014). *Student thinking and learning in science: Perspectives on the nature and development of learners' ideas.* Routledge.

Vygotsky, L. S (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Chapter 5

Linking the Correct Answer in Mathematical Problem Solving to a Veracious Statement in Everyday Conversation

Sergei Abramovich*
State University of New York, Potsdam, NY, USA

Abstract

The chapter reflects on the author's experience teaching asynchronously required mathematics content and methods course to elementary teacher candidates. The course includes weekly discussion forums two of which are "The importance of correct answers" and "Fostering accuracy in making claims". Responses provided by the candidates over three consecutive semesters were analyzed to suggest that learning mathematics does nurture skills to not say things that are untrue, both in professional and everyday conversations. The chapter shares many comments by future teachers of elementary mathematics that demonstrate their appreciation of logic, honestly, intellectual integrity, and like traits stemming from the course content. Several classroom episodes included in the chapter are connected to relevant historical and contemporary perspectives from mathematics, philosophy, and social sciences.

Keywords: mathematics, teacher education, problem solving, asynchronous teaching, habits of mind, everyday conversation

* Corresponding Author's Email: abramovs@potsdam.edu

In: Progress in Education. Volume 91
Editor: Roberta V. Nata
ISBN: 979-8-90134-375-3

Introduction

This chapter was written to share the author's experience of teaching mathematics asynchronously to future teachers of elementary (K-6) school mathematics of United States and Canada. This new type of teaching modality is becoming more and more practical, especially after the plight of COVID 19 that has affected education and its long-established tradition of teaching face-to-face. At the same time, a review of online teacher education programs in the United States (Huang & Manouchehri, 2019) indicated substantial challenges in teaching mathematics education courses online. One of the challenges is the lack of conventional teacher-student interaction that makes it difficult to motivate teacher candidates to learn mathematics. Fortunately for the author and hopefully for his students (many of whom are residents of the Canadian province of Ontario), asynchronous teaching includes teacher candidates' participation in discussion forums embedded into a Learning Management System (Brightspace) used by the university that employs the author. The discussion topics, offered to students each week of a semester-long (3 credit hours) course, are diverse and designed to address a variety of issues associated with contemporary mathematics education. Two topics included in the discussion forum grew out of the author's belief and life experience that mathematics with its educational emphasis on the diversity of conceptual thinking in problem solving and its eventual convergence to the right answer enables one to pay more attention to the veracity of statements made in a non-mathematical everyday conversation. Put another way, the chapter was written to suggest that mathematics education prepares learners of the subject matter, teacher candidates included, to be effective and authentic communicators in a variety of real-life contexts.

The chapter's content stems from two discussion forums titled "The Importance of Correct Answers" and "Fostering Accuracy in Making Claims" in which some 320 teacher candidates enrolled in the graduate level course "Elementary Mathematics Content and Methods" have participated over three consecutive semesters (fall 54, spring and fall 2025). The candidates' responses to the prompts "Share your perspective on the importance of correct answers in mathematical problem solving" and "Offer your perspective on the assertion that learning mathematics fosters our ability to not say things that are untrue" were reviewed and analyzed. The chapter includes selective responses that are typical for future teachers of elementary school mathematics. Figure 1 shows the results collected over three consecutive semesters from students responding to the prompts of the above two forums. One can see that the smallest and the largest number of words are 33 and 502, respectively, but the average number of words is pretty consistent across three semesters and two forums being in the range [145, 184]. There was no requirement for the minimum number of words used in the discussion forums as such a requirement, while would raise the minimum, would also decrease the maximum. For example, it appears that setting the minimum at 145 words would make the average word-range smaller.

Importnance of Correct Answers in Problem Solving			Making Veracious Statement in Everyday Conversartion		
words min	words max	words average	words min	words max	words average
33	326	184	53	341	165
60	391	145	63	502	148
68	454	174	36	357	164

Figure 1. Min, max, and average number of words as responses to forums by 320 students.

Logic, Mathematics, and Two Classroom Episodes

Mathematical reasoning honors logic and justification, both formal and informal. According to Walton (1992), ancient Greeks "saw logic as the art or skill of the use of reasoning in conversation or dialogue … to recognize and deal with fallacies … seen by Aristotle as common types of sophisticated deceptions or tricky tactics that can be used unfairly to get the best of an opponent in argument" (p. 1). Only beginning from the 20th century, due to advances in logic as part of mathematics (Russell, 1913; Kac & Ulam, 1968; Quine, 1981), psychology (Piaget & Inhelder, 1958), and, more recently, mathematics education (Durand-Guerrier, 2020), scholars returned to the ancient idea of studying everyday conversation through the lens of fallacies (Copi, 1953; Hamblin, 1970; Kahane, 1982; Tindale, 2007; Nelson, 2016). Whereas fallacies were described by Aristotle (1955) as "sophistical refutations", which mean incorrect methods of argument (Copi, 1953) or unfair verbal techniques used in conversation to win an argument (Walton, 1992), mathematics provides its learners with fair advantage to be more precise and logical when making deductively tenable arguments. That is, mathematics enriches people with meticulous skills to say things that, while not being fallacious, are difficult to defy because of being logical and rigorously justifiable.

As a simple example of fair advantage that mathematics provides, consider the game played by several kids in which the winner is the first who names the largest two-digit number by starting with the number 3 and continuously adding the number 5. A child recognized that through this addition process, the numbers have either 3 or 8 as the last digit and immediately won the game by naming 98 as the target number without actually doing multiple additions. In this game, the child used mathematics to win it, and the technique, while being sophisticated and somewhat tricky from a non-mathematical point of view, was fair and not deceptive. Furthermore, same technique can be used to defy the answer 99 because to reach 99, the largest two-digit number, by continuously adding 5, one has to start with 4. In this regard, one can recall the famous story of great Gauss, who as a student in grade school, was able, according to a legend (Dunham, 1991), to avoid the straightforward summation of the first 100 natural numbers – a task offered by his teacher – by recognizing that numbers equidistant from the beginning and the end of the sum add up to the same number.

Nonetheless, formal reasoning may be erroneous, especially in the case of mathematics where it is often more difficult to explain why something is incorrect rather than why something is correct. In the plain words of Hamblin (1970), describing Aristotle's primeval gist, "A fallacious argument … is one that *seems to be valid* but *is not so*" (p. 12, italics in the original). This description is consistent with the position of Kahane (1988) who did not see "the idea of deliberated deception as a condition of fallacy … [to allow for] the term to cover all *errors of reasoning*" (p. xv, italics in the original). With this in mind, as another invalid reasoning episode from an elementary classroom, consider a situation when a student "discovered" that the order relation between two common fractions does not change when each numerator and denominator is decreased by one. The following two pairs of common fractions, (3/7, 5/9) and (8/11, 6/7), were provided as justification. Indeed, 3/7 < 5/9, 2/6 < 4/8, and 8/11 < 6/7, 7/10 < 5/6 are all true inequalities. However, the "discovery" by the student, while, in general, is incorrect and, apparently, was not presented with an intent to deceive the teacher, is also difficult to defy because a counterexample, in the form of a pair of common fractions, (7/17, 5/12), is not easy to find (although it is not unique). One can check to see that whereas 7/17 < 5/12 the decrease mentioned by the student yields 6/16 > 4/11. Using other pairs of common fractions, one can also check to see that the student discovered not an atypical (but quite a convincing) case thereby prompting "hasty conclusion" (Kahane, 1988, p. 77).

Unlike formal proof that shows the correctness of a statement, a counterexample is a special mathematical tool used by mathematicians to show that a statement is untrue. In other words, both proof and counterexample are techniques used to decide what is true and what is untrue. Teacher candidates, the author's students, are familiar with the use of counterexamples as one of the candidates noted, "*Students can discover that their claim is untrue using counterexamples*." Most likely, though, by counterexample the candidate meant "clear exception" (Walton, 1992, p. 78) like in the case of presenting the number 2 as the only prime which is not odd.

One can see that in the first classroom episode, the number 99 was defied not through a counterexample but by demonstrating how it can be reached, and in the second episode, the student's "discovery" was defied by means of a counterexample. Much more complicated is the statement that every even number beginning from four is a sum of two prime numbers (Goldbach, 1742). The statement, named after a German mathematician Christian Goldbach (1690–1764) and known as Goldbach conjecture, remains neither proved nor disproved. Both techniques – confirmation through proof and refutation through counterexample – help one to develop a mentality (alternatively, habits of mind) of not saying things that are untrue.

Developing Habits of Mind through the Study of Mathematics

The ideas of redesigning mathematics curricula to encourage mathematical experiments as means of students' conceptual development were initiated more than half a century ago encouraging a forward-looking perspective on mathematics education as an "experimental science" (Begle, 1969, p. 242) to enable teachers of the subject matter to teach by "illustrating the nature of analogy between physical happenings in different engineering areas" (Richards, 1973, p. 386) and, in the age of technology, leading to the advances of experimental mathematics (Arnold, 2005; Borwein, 2005). Supported by classic conceptions of integrating a "mechanical method" (Archimedes, 1912, p. 13) into the study of geometry and 'quasi-experiments' (Euler, 1849) into the study of number theory as vital preludes to formal demonstration, these ideas gave birth to the modern standpoint about the development in high school students the habits of mind of mathematicians in order "to close the gap between what the users and makers of mathematics *do* and what they *say*

(Cuoco et al., 1996, p. 376, italics in the original). Over the last three decades, the ‘habits of mind’ standpoint was expanded to include earlier stages of learning mathematics as well as teacher preparation programs to focus on trial and error in problem solving, hands-on and computational experiments, multiple solution strategies, more than one correct answer in open-ended queries, and research-like activities (Clements, 1999; Goldenberg et al., 2003; Leikin & Dinur, 2007; Mark et al., 2010; Abramovich et al., 2012; Bullock, 2016; Carreira at al. 2016; Abramovich, 2020; Abramovich & Freiman, 2023). All these ideas were integrated into the course through problem-solving assignments and online exchange of views through weekly discussion forums. The two such forums considered in this chapter were scheduled back-to-back in the middle of a semester to allow teacher candidates to respond to the corresponding prompts with experience and knowledge of the modern-day problem-solving approaches to school mathematics already gained from the course. To a certain extent, the exchange of views on a particular topic can be described through the lens of triangulation, the method of integrating alternative epistemologies into the study of the same object typically used by social science researchers (Webb et al., 1966; Denzin, 2007) as a means of bypassing bias and ensuring rigor in the discipline’s conclusions. Recently the notion of triangulation was extended to mathematics education as more than one way of solving a problem using multiple digital tools (Abramovich, 2022, 2026).

Intellectual Conflicts as Means of Leaning Mathematics

As mentioned by Vinner & Tall (1982) in the context of teaching mathematics at the tertiary level, when making statements people “believe that these statements relate to a certain reality” (p. 752) and for students who are not mathematical majors “intellectual conflicts are not productive” (p. 755). The first quote points at the importance of using genuine problems in teaching mathematics to make it applicable to real life (Lyness, 1968). The second quote, however, may not be taken to mean that intellectual conflicts should always be avoided when teaching the subject matter, especially at the primary level. In fact, as many teacher candidates reported through the forum discussions, learning mathematics motivates “*questioning assumptions*”, something that seems to be the core of an intellectual conflict in the mind of a child. Indeed, two types of questions that children ask, requesting information and seeking explanation (Isaacs, 1930), are indicative of natural skills to seek clarification to a statement that is either doubtful or creates a cognitive incoherence. That is, in the words of a teacher candidate, "*instead of just accepting information at face value, you learn to double-check facts*". For example, whereas the equality 5 + 6 = 11 could be presented in a kindergarten classroom as an obvious fact, it might appear counterintuitive and thus create an intellectual conflict when one understands the meaning of addition but is not familiar with the concept of base ten. Indeed, both addends in the left-hand side of the above equality are greater than each of the two digits that comprise their sum. This is a reliable reality of the process of augmentation of quantities for young learners of the arithmetic of natural numbers. Another basic assumption of the arithmetic of early grades is the relationship between the minuend and the subtrahend. It is through questioning the assumption that the former may not be smaller than the latter that one resolves the emerging intellectual conflict, thereby enabling arithmetic being extended to include and, as a way of embracing real life, to give meaning to negative numbers. Mathematics, a subject matter that is full of such intellectual conflicts, helps already young children to recognize inconsistences in conversations outside the classroom and develop a mindset towards intellectual understanding of what is true and why (Piaget, 1959).

What Teacher Candidates Say

Sharing their views regarding the assertion that learning mathematics helps not saying things that are untrue, teacher candidates by and large were united when talking about mathematics helping them to create habits of "*logical analysis and careful expres*sion", "*truth-seeking*", "*honesty validation and logical scrutiny*", "*being careful*", "*reflecting on assumptions*", "*verifying information to ensure that statements are accurate and logical*", "*enhancing honestly and clear communication*", "*accuracy and logical consistency*", and "*thinking carefully and checking facts before making claims*". Using the words logical analysis by a candidate most likely means "interpretation, evaluation and construction of arguments and argumentation used in natural language" (Blair & Johnson, 1987, p. 148), alternatively, informal logic used "as a tool … of *everyday* argumentation" (Walton, 1992, p. 32, italics added). Likewise, many teacher candidates confirmed that through the study of mathematics they developed mindset that "*values accuracy*", "*prioritizes truthfulness and intellectual integrity*", "*helps to become more thoughtful and truthful in how to communicate and make decisions*", "*values truth based on proof rather than* *belief*", "*shapes how we approach truth in general*", and "*make claims only with evidence*".

In the words of one teacher candidate, learning mathematics develops "*disciplined way of thinking and nurtures intellectual honesty – admitting when we don't know something or when our reasoning doesn't hold up*". This comment about intellectual honesty deserves special consideration. Not all people readily admit not knowing something. It appears that the more one knows, the more likely one would say "I don't know". The inverse statement appears being true as well: the less one knows, the less likely one would say "I don't know". Some teachers might insist on students' solving problems only using the method they demonstrated in class. This position is likely due to reluctance to admit not knowing other ways as a teacher has to decide on the correctness of alternatives.

Nonetheless, when the course work emphasizes multiple solution strategies in problem solving, teacher candidates' responses to prompts of the corresponding forums indicate their conceptually oriented appreciation of the interplay between the right answer and solution that led to this answer. For example, one teacher candidate left the following comment in the forum regarding the importance of correct answers: "*I think correct answers in math are important, but sometimes the process is even more important.*" Such focus on problem solving process is commendable for a future teacher of elementary mathematics because learning the subject matter stems from this process. Another teacher candidate left a similar response: "*The correct answer is only a piece of the puzzle with other pieces including application and understanding*." In the next section, a puzzle would be presented through three unpretentious algorithms that students can use when solving a primary school problem in the context of using linking cubes in different colors. Despite leading to the same answer, the applicability of the algorithms will require recourse to special cases as means of mathematical understanding and logical analysis.

A Special Case as a Means of Dealing with Alternative Solutions

Knowing the right answer when facing alternative solutions may not be enough as the right answer may result from incorrect reasoning. This position is confirmed by one of the author's students who believes that "*focusing solely* *on correct answers may overlook the richer learning that occurs through strategy selection, reasoning, conceptualization, and sense-making.* It is through "sense-making" when analyzing the correctness of "strategy selection" that incorrect reasoning can be detected. As an example, consider three solutions to the task of finding the number of ways to build a tower out of three cubes of different colors – red, blue, and green – assuming that colors may not be repeated in a tower. Note that all three solutions will have the same answer. In the first solution, a student says that regardless of which color cube we put at the top, there are two ways to permute the other two cubes, and because we have three colors, the number of ways is the sum $2 + 2 + 2 = 6$. In the second solution, a student says that there are three ways to select a cube for the bottom position, leaving two ways to select a cube for the middle position and one way to put the remaining cube at the top. Therefore, those ways have to be added to get $3 + 2 + 1 = 6$ ways. In the third solution, a student says that the top cube can be selected in three ways, and once the selection is made, there are only two ways to select the cube to put in the middle, and finally only one cube remains for the bottom position. The student draws a tree diagram and concludes that the number of ways to build a tower is the product $3 \times 2 \times 1 = 6$. As mentioned by a teacher candidate, "*Different students may approach a problem in different ways, and that's not only okay – it's great.*"

The greatness of "more than one way of doing things" (Advisory Committee on Mathematics Education, 2007, p. 18) by the learners of mathematics points to the need to verify the correctness of each of the three solutions (algorithms) with the same, apparently correct, answer. One way of doing it is through verifying reasoning involved for a special case of two-cube tower with two colors. In the first solution, when only two colors may be on the top of the bottom cube, the sum $1 + 1 = 2$ gives an obviously (i.e., experimentally confirmed) right answer. In the second solution, there are two ways to select the top cube leaving only one option to the bottom cube yielding the sum $2 + 1 = 3$ which does not confirm the right answer. In the third solution, the right answer is confirmed as $2 \times 1 = 2$. Both the first and the third counting algorithms can be applied to the case of a four-cube tower and four colors demonstrating how classroom analysis of solution strategies, using the above words of a teacher candidate, provide context for *application* to problem solving and means of *understanding* the power of verification using reduction to special cases.

At a more complex level, one can note that whereas the $2 + 2 + 2$ algorithm is the legitimate application of the Rule of Sum used in combinatorics for counting mutually independent choices, the $3 + 2 + 1$

algorithm represents an erroneous use of the Rule of Sum as the second addend, 2, is not independent on the first one, 3, and the value of the third addend, 1, is determined by the first two addends, 3 and 2. A tree diagram drawn by one of the students, visually representing the Rule of Product, used in combinatorics for counting dependent choices, can show that the numbers 3, 2, and 1 must be multiplied as it was done correctly in the third solution. Put another way, the Rule of Sum in the second solution was a "sophistical refutation" (Aristotle, 1955) – an unfair use of this rule to get the right answer.

From a Forum to an Early Elementary Classroom

One teacher candidate left the following educationally notable comment as a response to the forum about developing truthfulness in the context of doing proof: "*We cannot simply say that 1 + 2 + 3 is the same as 4 + 2 if we do not have proof that this is true*." This comment not only reminds us of arithmetical operations demonstrated as problem-solving algorithms in the previous section. It deserves special attention as having important implications for teachers' work with students at the kindergarten level.

Figure 2 shows: (i) how the image of 2 + 4 can be re-arranged to have the image of 1 + 2 + 3; (ii) how both images are modified to have either two identical columns or three identical rows (a precursor to multiplication as repeated addition); (iii) alternative modifications of case (ii) as precursors to commutative property of addition, lateral and rotational symmetries, triangular numbers as partial sums of natural numbers; (iv) proving the sameness through one-to-one correspondence, and (v) finding the smallest number of ways the counters forming triangle can be repositioned to form two identical columns (or rows) and vice versa.

Such classroom elaboration of the comment by a teacher candidate regarding the need for proof in claiming the sameness of two sums shows what it might mean for a teacher to "orient his work not on yesterday's development in the child but on tomorrow's" (Vygotsky, 1987, p. 211). The suggested activity, representing ideas for a simple, yet conceptually rich hands-on mathematical experiment, can be supported historically by referencing Kepler's "without proper experiments I conclude nothing" (cited in Burtt, 1925, p. 50). It is also a reflection on didactical phenomenology of mathematics as "a way to show the teacher the places where the learner might step into the learning process of mankind" (Freudenthal, 1983, p. ix). In that way, as prompted by the teacher candidate within a forum, concrete materials as means of proof can be "used in the context of educational tasks to actively engage children's thinking with teacher guidance" (Clements, 1999, p. 56). As far as the idea of the 'habits of mind' is concerned, according to Goldenberg et al. (2003), "teaching these ways of thinking, which we refer to as *mathematical habits of mind*, is a vital part of mathematical instruction at every level" (p. 2, italics in the original). As an aside, in agreement with the view that argument and inference are two distinct tools of reasoning (Johnson, 1987), note that whereas demonstrating one-to-one correspondence between the images of the sums 1 + 2 + 3 and 2 + 4 is an argument and turning the ladder into the triangle (Figure 2) is an inference, both tools allow one to prove the equality mentioned in the above comment by a teacher candidate.

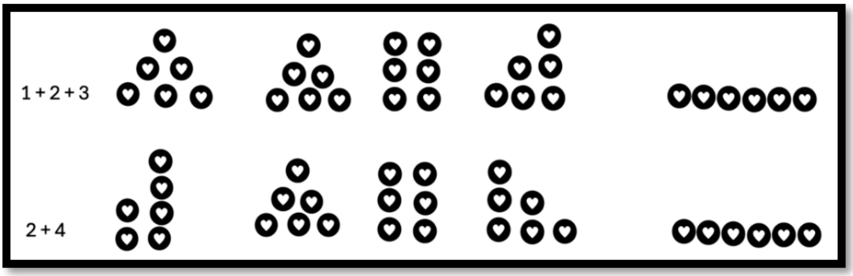

Figure 2. Different ways of proving that 1 + 2 + 3 is equal to 2 + 4.

Likewise, Figure 3 shows how a kindergarten teacher can use visualization to ask children to decide on the sameness of different representations of the two sums. The first row intends to show the same number of blocks each sum is comprised with. The second row aims at the verification of the concept of the conservation of number (Piaget, 1952) by students when the far-right tower from the first row turns into a train. The third row shows that depending on the unit of which the number is built, numbers can represent quantities of different kinds, something that is more complicated than the conservation of number; that is, in the left-hand side of the third row the numbers describing the corresponding images in the first and the second rows are not conserved. The fourth row shows conservation of numbers when the units, of which they were previously built, have changed. Under the guidance of a teacher, such activities would build the needed 'habits of mind', considering that, as mentioned by Estes (1956), neither age nor economic levels of students are necessarily factors in acquisition of mathematical ideas through visualization. That is, just as the fraction 1/2 has different meanings depending on which unit it is a part, the number 2 may also have different meanings depending on the unit it represents. It appears that through such conceptual teaching of mathematics already at the kindergarten level, aimed at the recognition of the interplay between quantities and their respective units, one's ability to not say things that are not true develops.

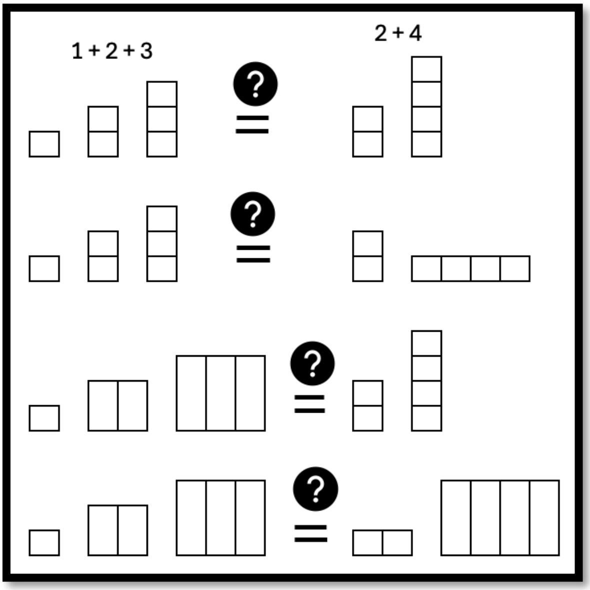

Figure 3. Do all images prove the sameness of 1 + 2 + 3 and 2 + 4?

Another comment by a teacher candidate, "*In the primary grades, I see this [mathematical habit of mind] when students talk through how they solved a problem or use models to show their thinking*", points at the call for proving the relation 1 + 2 + 3 = 2 + 4 as part of hands-on mathematics curriculum for young children aimed at not doing procedural arithmetic to get 'the right answer' but rather using conceptual understanding of what the sums in both sides of the relation might represent in real life. It is such understanding that develops mathematical 'habits of mind' already at an early age and creates foundation of enduring skills and divergent thinking (Bai, et al., 2024), something that another teacher candidate might have had in mind when noting: "*As an educator, I see how even young children begin developing awareness that their reasoning must make sense and align with what's true within the* *problem. A deeper respect for truth shows up not just in numbers but in the way we think, speak, and justify our ideas in everyday life*". The candidate's connection of respect for truth in numbers to justification of ideas related to everyday life supports the author's belief that learning mathematics enables one to pay more attention to the veracity of statements made in casual conversations.

Different Approaches to Dealing with Errors in Mathematics Education

Rushton (2018) found that teaching error analysis in a mathematics classroom creates in students a lasting habit of recognizing, as Common Core State Standards (2010) put it, the value of "plausible arguments … and – if there is a flaw in an argument – explain what it is" (p. 7). Alvidrez et al. (2024) suggested that there are two schools of thought among teachers about mathematical errors – errors "as resources for advancing understanding … [and] errors as deficiencies, to be avoided if possible and fixed quickly if not" (pp. 113, 114). The latter school of thought is recognized in the Standards for Preparing Teachers of Mathematics (Association of Mathematics Teacher Educators, 2017) where one can find a note that sometimes, mathematics teachers when facing students whose thinking is different from what is expected (including erroneous thinking), "may inadvertently seek to remedy those differences rather than seeing them as strength and resources upon which to build" (p. 22). Indeed, young children are very smart, each in their own way and thus "none of the child's reactions must go to waste" (Vygotsky, 1997, p. 347). As a teacher candidate put it, "*without wrong answers growth and further research and study would not continue*". Therefore, especially in education, an incorrect answer may be used as a thinking device (Mehan, 1979; Powell et al., 2012) and, as other teacher candidates noted, "*mistakes are important to aid in development*" and "*being wrong is part of getting it right*".

In the age of technology, errors may be of a computational nature as the accuracy in using a particular code even within a single digital tool depends on the accuracy of notation. A slight change in notation may be very consequential yielding inadvertent results. As noted by Conole & Dyke (2004), an easy access to the Internet requires one to learn how to manage the abundance of (sometimes incorrect) information provided, especially when artificial intelligence (AI) is involved. As mentioned by a teacher candidate, "*I am not sure if I'm taking this too far, but the forum makes me think about AI debate. AI lives through intelligence, reasoning, and critical thinking, but does it have a moral compass*?". Writing about uses of AI in mathematics education, Opesemowo & Ndlovu (2024) warned that overdependence on AI with little (if any) conceptual understanding may lead to the loss of critical thinking and the neglect of algorithmic biases. The author believes that unfortunate errors in students' (and others) uses of AI are not due to the absence of the tool's "*moral compass*" but rather due to the lack of conceptual understanding and/or critical thinking by its users. In particular, a recourse to computational triangulation (Abramovich, 2026), as using more than one digital tool in problem-solving as an effective oversite of computational errors can protect problem solvers from the algorithmic biases of AI. To a certain extent, an algorithmic bias of AI coupled with the users undue reliance on machine learning as a substitute for their own knowledge and experience can lead to the modern day type of "sophistical refutation" (Aristotle, 1955) when, as was observed by the author in a teacher candidate's assignment, a machine-generated 4 by 12 rectangle on a geoboard was in fact a 3 by 11 rectangle. Because on a geoboard the latter rectangle is associated with 4 and 12 pegs, it might be that AI was asked by the candidate to construct a former rectangle by connecting those pegs. Consequently, an error was missed visually because the candidate's loss of critical thinking was due to their overdependence on AI.

Learning to Recognize and Avoid Incorrect Answers Using Common Sense

As was mentioned in the introduction, the forum about mathematics as a milieu for developing skills in making veracious statements in everyday conversation was a follow-up to a related forum about the importance of correct answers in problem solving. In the latter forum teacher candidates also expressed their opinion on the issue of correct answers. Most of the respondents, admitting the importance of providing a mathematical question with a correct answer, argued that the process through which the answer was obtained is not less important than the answer itself. Among the responses are: "*Even an incorrect answer can lead to meaningful learning if you take the time to reflect on it*"; "*Focusing solely on correct answers overlook the richer* *learning that occurs through strategy selection, reasoning, conceptualization, and sense-making*", "*When teachers encourage students to explain their thinking, justify their reasoning, and explore multiple methods to reach a solution, they promote higher-order thinking and resilience in learning*", "*Correct answers are based on proven and evidence-based processes rather than formed by perspective or opinion*", "*As educators, it's our job to show students strategies to find the correct answers, but we also have to allow for diverse thinking*", "*The importance of correct answers when students' efforts are valued they feel more confident tackling complex problems in their own way*".

The analysis of the above comments suggests that reflecting on an incorrect answer fosters the ability of thinking about cognitive rectification when the conversational statement that first comes to mind is untrue. This kind of reflection is virtually unconscious and almost instantaneous, yet, without experience of reflecting on an incorrect answer, the need for such a quick self-checking would not come to mind. However, as the discussion of three alternative solutions leading to the same answer demonstrated, reflection on the correct answer, as a teacher candidate noted, is not enough. Rather, problem-solving reflection should not overlook the "*strategy selection*" that led to the answer. Furthermore, the discussion's second strategy of *adding* the number of ways to build a three-cube tower, in order to be understood as a deficient one, required a counterexample in the form of recourse to a special case; alternatively, revealing the Rule of Sum as a strategy "that *seems to be valid* but *is not so*" (Hamblin, 1970, p. 12, italics in the original). Revealing to teacher candidates such interplay between the strategy and the answer is not to discourage problem solvers from "*tackling complex problems in their own way*" but rather to show how complexity of mathematics can be demonstrated through a simple hands-on task. This demonstration does promote higher order thinking without affecting negatively "*resilience in learning*".

Focusing on a problem-solving process is important provided that the very process is correct. For example, in the course, the candidates were asked to represent the fraction 3/5 as a sum of two unit fractions, such representation known as an Egyptian fraction. The question was put in the real-life context of fair sharing 3 identical pizzas among 5 children so that the equality 3/5 = 1/2 + 1/10 can be interpreted as giving each child two slices measured by 1/2 of a pizza and by 1/10 of a pizza (Figure 4, the top part). But when asked about the total number of slices obtained through this division, some teacher candidates added the denominators of the two fractions saying that the total number of slices is twelve without realizing that counting the slices happens at the level of common sense outside the domain of fractions. And while 12 is not much different from the correct answer 10, in the case of dividing 5 pizzas among 6 people, the representation 5/6 = 1/ 2 + 1/3 (Figure 4, the bottom part) would have a larger difference between incorrect, 5, and correct, 12, answers. But the incorrect answer, 5, is smaller than the number of children involved. Such recourse to common sense, that develops in early age through social interaction with peers and more knowledgeable others (Vygotsky, 1987), allows for recognition of inconsistency in argument. The inconsistency is due to one's understanding of real-life context of pizza sharing that is an indicator of using informal logic and everyday cognition as tools of making a true argument. The presence of such examples when the answer contradicts common sense is important for developing skills of making veracious statements in everyday conversations that are often structured by informal logic.

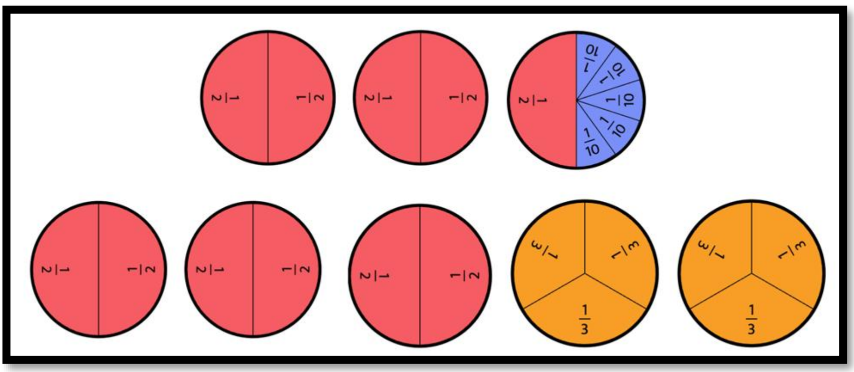

Figure 4. Top: $\frac{5}{6} = \frac{1}{2} + \frac{1}{3}$ as 5 pizzas for 6; Bottom: $\frac{3}{5} = \frac{1}{2} + \frac{1}{10}$ as 3 pizzas for 5.

Conclusion

Just as common sense develops in early ages through social interaction with peers and adults, conversational skills develop within a cultural milieu of K-12 education. Elementary teacher candidates, prompted to reflect on the development of their conversational skills through the learning of mathematics, described mathematical education as an intellectual environment within which logical reasoning and justification of arguments, as agents of problem solving converging to the right answer, nurture veracity.

The chapter provided a number of examples from school mathematics – the subject matter studied by all future schoolteachers – enabling the creation of mathematical habits of mind, drawing on cognitive conflicts and real-life situations as means of conceptual development, making use of common sense as a reasoning instrument, and resorting to special cases as counterexamples of seemingly valid solution strategies. Such studies, as was confirmed by teacher candidates, affect one's probity when making claims in everyday conversations. The chapter was also supported by referencing historical and contemporary perspectives from mathematics, philosophy, psychology, and sociology to edify the modern-day classroom episodes related to visualization, pattern recognition, hands-on experimentation with concrete materials, and the growth of divergent thinking, all things oriented on tomorrow's development of the learners. In particular, an incorrect use of the Rule of Sum and students' undue reliance on machine learning have been described as the modern-day "sophistical refutations" leading, respectively, to correct answer with incorrect solution and erroneous geometric construction on a digital geoboard.

To conclude, the value of discussion forums as part of an online mathematics education course that strives to balance its pros and cons should be acknowledged. The forums make it possible for instructors to integrate teaching and scholarships. Whereas online pedagogy endures the lack of traditional student-teacher interaction, appropriately chosen topics for online discussions become enlightening magnets enticing future teachers' participation in noninvasive educational settings that immune from the time constraints (as data in Figure 1 might indicate) of the curriculum to be delivered. By analyzing two discussion forums, the chapter visibly demonstrated that learning mathematics, with its reverence for correct answer, enriches teachers, as major custodians of knowledge, with useful skills and strategic disposition of not saying things that are untrue in the classroom and beyond.

References

Abramovich, S. (2020). A note on teaching mathematics to elementary teacher candidates: Cookies, creatures, Fibonacci numbers and spreadsheets. *Open Mathematical Education Notes*, *10*(1), 15-32.

Abramovich, S. (2022). Advancing the concept of triangulation from the social sciences research to mathematics education. *Advances in Educational Research and Evaluation*, *3*(1), 201-217. DOI 10.25082/AERE.2022.01.002

Abramovich, S. (2026). *Computational Triangulation in Mathematics Teacher Education*. Nova Science Publishers.

Abramovich, S., Easton, J., Hayes, V. O. (2012). Parallel structures of computer-assisted signature pedagogy: the case of integrated spreadsheets, *Computers in the Schools*, *29*(1–2), pp. 174–190.

Abramovich, S., and Freiman, V. (2023). Fostering collateral creativity through teaching school mathematics with technology: what do teachers need to know ? *International journal of mathematical education in science and technology*, *54*(10), 2217-2242.

Advisory Committee on Mathematics Education. (2007). *Mathematical Needs of 14–19 Pathways*. The Royal Society.

Alvidrez, M., Louie, M., & Tchoshanov, M. (2024). From mistakes, we learn? Mathematics teachers' epistemological and positional framing of mistakes. *Journal of Mathematics Teacher Education*, *27*(1), 111–136.

Archimedes. (1912). *The Method of Archimedes*. Edited by T. L. Heath. Cambridge University Press.

Aristotle. (1955). *On Sophistical Refutations*. E. S. Forster (trans.), Loeb Classical Library Edition. Harvard University Press.

Arnold, V. I. (2005). *Experimental Mathematics* (in Russian). Fazis: Moscow, RF.

Association of Mathematics Teacher Educators. (2017). Standards for Preparing Teachers of Mathematics. Available online at: https://www.amte.net/.

Bai, H., Mulder, H., Moerbeek, M., Leseman, P. P. M., & Kroesbergen, E. H. (2024). The development of divergent thinking in 4- to 6-year-old children. *Creativity Research Journal*, *36*(4), 656–674.

Begle, E. (1969). The role of research in the improvement of mathematics education. *Educational Studies in Mathematics*, *2*(2/3), 232–244.

Blair, J. A., & Johnson, R. H. (1987). The current state of informal logic. *Informal Logic*, *9*(2), 147–151.

Borwein, J. M. (2005). The experimental mathematics: The pleasure of discovery and the role of proof. *International Journal of Computers for Mathematical Learning, 10*, 75–108.

Bullock, S. M. (2016). Teacher candidates as researchers. In J. Loughran & M. L. Hamilton (Eds.), *International Handbook of Teacher Education*, vol. 2 (pp. 379–403). Springer.

Burtt, E. A. (1925). *The Metaphysical Foundation of Modern Physical Science*. Kegan Paul.

Carreira, S., Jones, K., Amado, N., Jacinto, H., & Nobre, S. (2016). *Youngsters Solving Mathematical Problems with Technology.* Springer.

Clements, D. H. (1999). 'Concrete' manipulatives, concrete ideas, *Contemporary Issues in Early Childhood Education*, *1*(1), pp. 45–60.

Common Core State Standards. (2010). *Common Core State Standards Initiative: Preparing America's Students for College and Career*. [On-line materials]. Available at: http://www.corestandars.org.

Conole, G., & Dyke, M. (2004). What are the affordances of information and communication technologies? *ALT-J: Research in Learning Technology*, *12*(3), 113–124.

Copi, I. (1953). *Introduction to Logic*. Macmillan.

Cuoco, A. , Goldenberg, E. P., & Mark, J. (1996). Habits of mind: An organizing principle for a mathematics curriculum. *Journal of Mathematical Behavior*, *15* (4), 375–402.
Denzin, N. K. (2007). Triangulation. In G. Ritzer (Ed.), *The Blackwell Encyclopedia of Sociology*, vol. X (pp. 5075–5080). Blackwell Publishing.
Dunham, W. (1991). *Journey through genius: The great theorems of mathematics*. Pantheon Books.
Durand-Guerrier, V. (2020). Logic in mathematics education. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education* (pp. 478–481). Springer.
Estes, B. W. (1956). Some mathematical and logical concepts in children. *The Journal of Genetic Psychology*, *88*, 219–222.
Euler L. (1849) Regula facilis problemata Diophantea per numeros integros expedite resolvendi (An easy rule for Diophantine problems to be resolved expeditiously in integers). In Leonhardi Euleri Commentationes Arithmeticae (in Latin, composed by P. H. Fuss and N. Fuss – great-grandchildren of Euler), (The Imperial St. Petersburg Academy of Sciences, St. Petersburg, Russia) pp. 263-269.
Freudenthal, H. (1983). *Didactical Phenomenology of Mathematical Structures*. Reidel.
Goldbach, C. (1742). *Letter to L. Euler*, June 7, 1742.
Goldenberg, E. P., Shteingold, N., & Feurzeig, N. (2003). Mathematical habits of mind for young children. In F. K. Lester & R. I. Charles (Eds.), *Teaching Mathematics Through Problem Solving: Prekindergarten-Grade 6* (pp. 15–29). National Council of Teachers of Mathematics.
Hamblin, C. L. (1970). *Fallacies*. Methuen.
Huang, D., & Manouchehri, A. (2019). Online mathematics teacher education in the US: A status report. *Contemporary Issues in Technology and Teacher Education*, *19*(2), 171–194.
Isaacs, N. (1930). Children's why questions. In S. Isaacs, *Intellectual Growth in Young Children*. Routledge & Kegan Paul.
Johnson, R. H. (1987). Logic naturalized: Recovering a tradition. In F. H. van Eemeren, R. Grootendorst, J. A. Blair, & C. A. Willard (Eds.), *Argumentation: Across the Lines of Discipline* (pp. 48–56). Foris Publications.
Kac, M., & Ulam, S. M. (1968). *Mathematics and Logic*. Praeger.
Kahane, H. (1988). *Logic and Contemporary Rhetoric: The Use of Reason in Everyday Life*. Wadsworth.
Leikin, R., & Dinur, S. (2007). Teacher flexibility in mathematical discussion. *The Journal of Mathematical Behavior*, *26*(4), 328–347.
Lyness, R. C. (1968). Applied mathematics in English schools. *Educational Studies in Mathematics*, *1*(1/2), 98–104.
Mark, J., Cuoco, A., Goldenberg, E. P., & Sword, S. (2010). Contemporary curriculum issues: Developing mathematical habits of mind. *Mathematics Teaching in the Middle School*, *15*(9), 505–509.
Mehan, H. (1979). *Learning Lessons*. Harvard University Press.
Nelson, L. (2016). *A Theory of Philosophical Fallacies*. Springer.
Opesemowo, O. A. G., & Ndlovu, M. (2024). Artificial intelligence in mathematics education: The good, the bad, and the ugly. *Journal of Pedagogical Research*, *8*(3), 333–346.

Powell, J.C., Bernauer, J., & Agnihotri, V. (2012). Learning from Incorrect Answers on Multiple Choice Tests. In: Isaias, P., Ifenthaler, D., Sampson, D., & Spector, J. (Eds.), *Towards Learning and Instruction in Web 3*.0 (pp. 265–287). Springer.

Piaget, J. (1952). *The Child's Conception of Number*. Routledge & Kegan Paul.

Piaget, J. (1959). *The Language and Thought of the Child*. Routledge & Kegan Paul.

Piaget, J., & Inhelder, B. (1958). *The Growth of Logical Thinking*. Routledge & Kegan Paul.

Richards, E. J. (1973). Some thoughts on mathematics education. *International Journal of Mathematical Education in Science and Technology*, *4*(4), 377–395.

Rushton, S. J. (2018). Teaching and learning mathematics through error analysis. *Fields Mathematics Education Journal*, 3, 4. https://doi.org/10.1186/s40928-018-0009-y.

Russell, B. (1913). The philosophical importance of mathematical logic. *The Monist*, *23*(4), 481-493.

Quine, W. V. O. (1981). *Mathematical Logic*. Harvard University Press.

Tindale, C. W. (2007). *Fallacies and Argument Appraisal*. Cambridge University Press.

Vinner, S., & Tall, D. (1982). Existence statements and constructions in mathematics and some consequences to mathematics teaching. *The American Mathematical Monthly*, *89*(10), 752–756.

Vygotsky, L. S. (1997). *Educational Psychology*. CRC Press.

Vygotsky, L. S. (1987). Thinking and speech, In R. W. Rieber & A. S. Carton (Eds), *The Collected Works of L. S. Vygotsky*, vol. 1 (pp. 39–285). Plenum Press

Walton, D. N. (1992). *Plausible Argument in Everyday Conversation*. SUNY Press.

Webb, E. J., Campbell, D. T., Schwartz, R. D., Sechrest, L. (1966). *Unobtrusive Measures*. Rand McNally.

Chapter 6

Promoting Learning and Academic Success: A Principle-Based, Student-Centered Approach to Providing Feedback in College Writing Courses

Alicia J. Carthell*, EdD
Murray State University, Murray, Kentucky, USA

Abstract

Effective feedback promotes student academic success and learner agency and has many other positive effects in the higher education learning environment. This paper presents a student-centered approach to providing feedback on writing assignments in college writing courses. During dialogic draft review feedback sessions (DRFS), students receive customized, reliable, and specific feedback that encourages critical thinking, reflection, and self-evaluation, and they receive guidance on applying the feedback. During the DRFS, the goal is to learn about students' abilities, assess their needs, support comprehension, underpin motivation and confidence, enrich writing competencies, and promote learner agency. This paper contributes to the existing research on the significance and benefits of effective instructor feedback, outlining an approach centered on three key principles of feedback.

* Corresponding Author's Email: acarthell@murraystate.edu

In: Progress in Education. Volume 91
Editor: Roberta V. Nata
ISBN: 979-8-90134-375-3

Introduction

According to Hwang (2025), "Writing is a goal-oriented cognitive activity that requires metacognition and is essential for learners across all academic levels (p. 1). College writing classes are a central component of students' college experience, and the amount of writing that college students do is contingent upon their discipline (Cook, 2013). Writing is a key competency and essential tool for self-expression, learning, and academic success (Hopkins & Reid, 2024; Klein & Boscolo, 2016; Marlink & Wahleithner, 2011; McLeod, 2020). College classes assist students in developing strong writing skills that enable them to become better communicators in academia and in their professions. Whether in writing class or classes in their discipline, a central component of students' curriculum is writing in various genres and for various purposes. Students learn the importance of employing various communication approaches for diverse rhetorical situations that consider ethics and legal responsibilities (Anderson, 2018; Kolin, 2023; MacRae, 2019; McDonald et al., 2015). As students navigate academics and prepare for professional positions, writing for diverse audiences and for different purposes calls for them to build their portfolio of writing skills, which include critical thinking, critical writing, evaluation, audience awareness and analysis, rhetoric, clarity, conciseness, coherence, completeness, accuracy, style, tone, and layout/design. College writing classes span a variety of genres, including composition and research, business and technical writing, and journalism. In addition, some writing classes are tailored to students in certain majors, like classes that focus on writing in the social sciences, where students prepare for writing and research in their social science discipline. Irrespective of the type of writing class, engaging students in college writing courses can be a continual challenge for college faculty due to various factors, such as a lack of student-perceived relevance, rigid structures, and a lack of confidence in writing ability. Some ways that faculty engage students might include connecting assignments to real-world contexts, using multimodal formats, and incorporating service learning. One way to support students as they navigate college writing courses (some of which they are required to take) is constructive feedback on their writing. The impact of feedback, written and oral, has been studied from various vantage points (e.g., Bitchener & Storch, 2016; Boud & Molloy, 2013; Bush, 2021; Carhill-Poza et al., 2023; Ellis, 2010; Han, 2019; Lusher et al., 2021; Mulliner & Tucker, 2017; Rofiqoh, 2020; Tayebipour, 2019; Tehrani, 2018). Effective feedback is a central tenet of engaging with students, as it provides faculty with opportunities to provide customized learning that can enhance students' writing and overall academic success. Researchers have highlighted the many benefits of effective feedback and its role in student learning and academic success.

Literature on Feedback

For students in academia, feedback is an essential component of the writing process because one of the best ways to learn how to write is through continual evaluation (Tehrani, 2018). In education, Hattie and Timperley (2007) said that feedback is information given by an "agent concerning one's performance or understanding," and is a "consequence" of performance (p. 81). They asserted that feedback is an effective instructional tool for learning when blended with effective teaching. According to Rofiqoh (2020), teacher feedback (written or oral) plays an integral role in helping students improve their writing, and feedback helps guide students to other writing assignments; therefore, the feedback must be presented at the students' level so that they can comprehend it and engage with it. Rofiqoh also insisted that feedback should be clear, consequential, focused, and tailored to the students' knowledge to be effective. Brookhart (2008) stated that there is no one strategy that is compatible with all students. Ellis (2009) agreed and said there is no "corrective feedback recipe" (p. 106), and instructors should modify their feedback approaches to best meet the needs of the learners in specific contexts. With this in mind, educators should tailor feedback to students' needs as they learn about their strengths and areas of writing that require attention. Fathman and Whalley (1990), Williams and Jasmine (2003), and Furneaux et al. (2007) outlined the benefits of feedback and explained that feedback consists of form (grammar and mechanics) and content (organization, arrangement of ideas, and thought). Teacher feedback is designed to highlight both content and form in writing, thereby aiding student learning and enhancing their writing skills (Hyland, 2003).

Researchers also discuss self-regulation as a benefit of feedback. Nicol and Macfarlane-Dick (2006) interlinked effective feedback and the ability for students to self-regulate their own writing. The authors argued that self-regulation of one's writing is a critical feature of effective feedback, and students must be given opportunities to cultivate their ability to regulate their own learning throughout higher education to be prepared for learning in life. Hwang (2025) also insisted that feedback is a predictor of self-regulated learning and writing outcomes. In researching the impact of feedback perception and self-regulated writing ability, Hwang found that enhancing undergraduates' writing abilities and achievement necessitates that instructors provide constructive and positive feedback to improve students' perceptions of the feedback, underpin instructors' endeavors to increase self-efficacy, and bolster self-regulated learning. Feedback that is structured with actionable and timely evaluation is essential (Hwang, 2025). Similarly, Gombert et al. (2024) outlined the role of feedback in learning and the essentiality of constructive feedback, and they asserted that feedback must be highly informative to help students improve, as feedback is one of the most significant factors of educational instruction. Discussing the significance of feedback on learning and academic achievement, Gombert et al. believed that feedback must go beyond points and grades to allow for learning and achievement.

Emphasizing the significance of instructor feedback research, Cheng (2022) found that students indicated time and again that they preferred that feedback on their writing be direct and specific, because it helps them interpret the feedback. Other researchers also agreed that direct and specific feedback was necessary for students to comprehend the path forward. Mulliner and Tucker (2017) and Boud and Molloy (2013) found that students can feel discontented with the feedback they receive and the timeframe in which it is provided. According to the researchers, students found it challenging to understand and address feedback if it lacked detail and was ill-timed. Other researchers also touted the significance of timely feedback. Cárcamo (2020) claimed that timely feedback is a central resource of effective learning assessment. Nguyen (2019) agreed and said that "Instruction, practice, and feedback are timely when the instruction addresses learners' most relevant problems from their recent writing, the practice immediately follows the instruction, and the feedback is provided promptly after the practice" (p. 123). Feedback in the earliest stages of writing assignments can help facilitate the timely implementation of feedback and assistance.

Adding to the discussion about constructive, specific, and timely feedback, Mahboob (2015) insisted that effective feedback is purposeful, goal-oriented, and structured. Moreover, Mahboob advocated for cohesive and coherent feedback approaches, which are thoughtfully and wisely written content methods based on an evaluation of students' needs. The author stated that cohesive feedback encompasses the audience, organization, and the overarching goals of the writing, and coherent feedback is a framework that provides clear rationale and explicit feedback, helping students develop a profound understanding of their writing issues. Mahboob highlighted not only the importance of instructor feedback for students, but also the significance of supporting students' feedback comprehension and helping them to be self-regulating writers. Mahboob said that cohesive and coherent feedback helps students understand their writing strengths, weaknesses, and the issues they pose, as well as how to address and avoid these writing issues in the future. Instructors must provide feedback that is understandable and focused, so that students are not overwhelmed. Both Cook (2013) and Moore (2014) suggested that instructors not overwhelm students with excessive feedback, as it may ultimately be ineffective. Cook (2013) explained that feedback should be balanced enough for students to learn from and not be inundated. In the same vein, Moore (2014) suggested that when presented with immense feedback, students may feel overwhelmed and lose confidence. Moore suggested the "less is more" approach when giving feedback on academic writing.

Feedback on writing is not confined to correction; it helps us connect with students through customized, student-centered interactions and purposeful clarity, as we meet them where they are and help them grow into critical thinkers and proficient writers.

Draft Review Feedback Sessions

For all high-stakes assignments (HSAs), I hold DRFS where I observe students' progress, assess their abilities and comprehension, and offer feedback to support self-regulated learning. HSAs are summative assignments where students can demonstrate their understanding of course skills and concepts, and assignments that comprise a large percentage of the course grade (DePaul University, 2025). The HSAs reveal students' comprehension and ability to apply concepts and strategies covered in the curriculum, so I provide DRFS early and midway in the drafting phase to ascertain students' knowledge, understanding, and abilities, and to offer personalized instruction and assistance. Mao and Lee (2024) explained how feedback can cater to the individual needs of students.

> "To cater for individual learner needs, teachers may provide opportunities for individualized instruction – for example, take into account individuals' difficulties in understanding and using feedback, help students to set appropriate learning goals, and adjust feedback choices for individual learners. A personalized approach to feedback provision is likely to actively engage students in feedback practices,

> rendering feedback more student-centered and geared towards learners' specific needs." (Mao & Lee, 2024, p. 816)

These DRFS provide students with tailored feedback and assistance, enabling me to assess their skills and knowledge and identify areas that require support. During these sessions, I learn about comprehension and clarity issues, which help me modify future assignment criteria, lessons, discussions, and preceding exercises. The DRFS offer a safe space for students to ask questions and voice concerns without judgment, and provides me an opportunity to engage with them and build rapport. During DRFS, I provide targeted assistance for structural and contextual components of students' writing, in addition to content, style, tone, and adherence to the assignment criteria. Providing DRFS is central to my teaching because the feedback and assistance offer students a heightened understanding of curriculum concepts, principles, and writing strategies that aid in completing present and impending assignments.

I hold DRFS because of the numerous benefits I have observed for students in my writing classes. In effect, the DRFS I facilitate offer students the following benefits:

- provide opportunities for deeper comprehension of assignment criteria and expectations
- help them apply thoughtfully the knowledge, concepts, principles, and strategies from the curriculum
- help allay their fears, nervousness, or anxiety about asking questions or voicing concerns
- encourage self-reflection and self-assessment
- help boost motivation and confidence
- cultivate learner agency and self-regulation
- promote reflective practice
- encourage critical reading, thinking, and writing
- provide adapted assistance and actionable feedback
- promote feedback literacy (i.e., how to apply feedback)

The foregoing benefits for students also benefit my teaching, as I learn more about each student individually and can readily identify patterns or trends in writing strengths and areas that need improvement. Moreover, the information I observe from semester to semester helps inform my lectures, discussions, and activities that provide the foundation for the HSAs reviewed and assessed during the DRFS. The DRFS have been a learning experience for me and will continue to be a vital component of my teaching.

Purpose

The purpose of this paper is to discuss how the DRFS are aligned with components of Nicol and Macfarlane-Dick's (2006) feedback framework. The DRFS are centered on elements of Principles 3, 4, and 5 of Nicol and Macfarlane-Dick's (2006) "Seven Principles of Good Feedback Practice: Facilitation Self-Regulation" designed for feedback in higher education settings. The authors asserted that good feedback practice does the following:

- Principle 3: Delivers high quality information to students about their learning
- Principle 4: Encourages teacher and peer dialogue around learning
- Principle 5: Encourages positive motivational beliefs and self-esteem

(Nicol and Macfarlane-Dick, 2006, p. 205)

Promoting autonomy through reflective learning via feedback can have a positive impact on student academic success, and effective instructor feedback in college writing classes can help cultivate positive student-instructor relationships and contribute to positive learning environments for students. DRFS have helped me better engage with students and foster a supportive learning environment by working with them individually and in-person to assess their needs and address their writing concerns. This paper joins other research about the significance and benefits of effective instructor feedback in college writing courses, and discusses my strategic approach for cultivating a positive teaching and learning environment through actionable feedback and guidance for feedback application.

Principle 3: Delivers High Quality Information to Students About Learning

To make feedback actionable, I provide students with specific feedback on their writing, punctuation, language use, layout, citations, and tone of voice--

noting how the specific issue affects their overarching message to intended readers. Specificity provides them with the clarity they need to understand the issue(s). If I observe sentences with issues, I ask students to read sentences aloud. What I find is that once they read the sentences, they will often, with my guided feedback, identify the writing issue(s), evaluate their writing based on their message objectives, and then discuss with me possible ways for remediation. These moments of guided discovery help underpin reflection, critical thinking, autonomy, and high-quality learning. Providing students with feedback, comingled with guidance, and giving them the opportunity to reflect on and engage with the feedback, helps them become active participants in their own learning. Nicol and Macfarlane-Dick (2006) emphasized that "Good quality external feedback is information that helps students troubleshoot their own performance and self-correct: that is, it helps students take action to reduce the discrepancy between their intentions and the resulting effects" (p. 208). Based on my observations, when students obtain some emotional distance from their writing and read from an external perspective during the DRFS, they can ascertain if the intention of what they wrote aligns with their writing objectives and the assignment criteria. Nicol and Macfarlane-Dick (2006) stated that student reflection, which promotes self-correction and is connected to predefined criteria, is beneficial.

Furthermore, during DRFS, I encourage students to always consider their intended readers and adopt a reader-centered approach; however, this can be a challenge for students who often write primarily for themselves, resulting in self-focused writing. These DRFS help me to help them keep readers at the forefront of their writing, ensuring their writing is lucid, concise, accurate, specific, thorough, and coherent. I provide them with clear and actionable direction for creating reader-centered content, encouraging them to think critically and reflect on the assignment objectives and their overarching learning goals for the task at hand. I have found that helping students implement feedback in real time helps reduce ambiguity, models the revision process, and promotes both clarity and agency. This, in turn, often eases their anxiety, uncertainty, or nervousness and boosts their confidence for current and future assignments. Providing students with specific, clear, and actionable feedback in real-time, in-person, and individually supports their learning, as it encourages them to ask questions, voice concerns, and engage with the feedback. Researchers have emphasized the importance of feedback and argue that students must engage with it for it to be effective. Student engagement with the feedback is an essential component of effective feedback, and has been well researched (Han & Hyland, 2015; Mao & Lee, 2024; Nguyen, 2024;

Zhang & Hyland, 2022). According to Han and Hyland (2015), student engagement is what connects the feedback with learning outcomes.

DRFS are held well in advance of the assignment due date, so students have sufficient time to reflect and revise their drafts before submission. Nicol and Macfarlane-Dick (2006) outlined the need for feedback to be offered in a timely manner. Such feedback on assignments in progress affords students the opportunity to gain a deeper understanding of expectations and criteria during individualized instruction, and provides them with the chance to receive not only corrective feedback from me but also assistance in executing the feedback. The DRFS are student-centered in that the focus is not solely on giving corrective feedback and highlighting areas for growth; during the sessions, I also underscore and commend students' strengths and proficiencies. This student-centered approach stimulates both student engagement and learner agency (Mao & Lee, 2024), and provides individualized, high quality learning that is both cognitive and relational.

Principle 4: Encourages Teacher and Peer Dialogue Around Learning

While Nicol and Macfarlane-Dick (2006) present this principle in the context of both teachers and peers, my discussion of this principle is centered on the DRFS between students and me. During the DRFS, I make learning interactive by ensuring that feedback sessions are conversational—that they are dialogues rather than monologues. When students ask questions and voice concerns, I can support their learning by further explaining and clarifying concepts, principles, strategies, assignment criteria, and expectations. The DRFS are designed to be conversations—two-way exchanges where questions are invited, making learning a dialogic and collaborative process. In essence, during DRFS, which are immensely dialogic, students learn not only what to fix but also *why* remediation is essential and why it matters in terms of writing objectives, intended readers, or reader perceptions. In this reflective space, individualized teaching and learning become collaborative, as students and I discuss writing and rhetorical strategies and choices that influence how their writing might be received or perceived by readers. Moreover, this collaborative dialogue modality helps them move beyond subjective feelings and ponder how their writing and rhetorical choices impact the messaging. Furthermore, intentional and purposeful dialogue during DRFS also promotes collaborative problem-solving and knowledge building, helping students co-author their learning. Nicol and Macfarlane-Dick stated that feedback can be more effective and understood by students if it is part of a dialogue—a conversation not confined to information transmission. Rather, effective feedback entreats students to engage in discussion about the feedback and to receive a real-time response to writing issues. The authors noted, “Discussions with the teacher help students to develop their understanding of expectations and standards, to check out and correct misunderstandings, and to get an immediate response to difficulties” (Nicol & Macfarlane-Dick, 2006, p. 210).

Principle 5: Encourages Positive Motivational Beliefs and Self-Esteem

Nicol and Macfarlane-Dick (2006) wrote, “teachers can have a positive or negative effect on motivation and self-esteem” (p. 212). To encourage students’ motivation and positively influence their self-esteem during DRFS, I acknowledge the positives of their draft—recognizing their progress and praising their strengths and proficiencies. “Praising effort and strategic behaviors, and focusing students through feedback on learning goals, leads to higher achievement than praising ability or intelligence” (Nicol & Macfarlane-Dick, 2006, p. 212). To help motivate students and have a positive effect on their self-esteem, I am mindful of my language use and employ questions that foster reflective practice, critical thinking, and questioning. I filter feedback through a questioning framework—using inquiry and guidance to help them discover a way or multiple ways to resolve writing issues in their draft, and then apply the feedback through revision. One example question I use when asking students to assess their writing is “*how can we say this to eliminate the abrasive tone so that readers won’t feel attacked; after all, our goal is to have a conversation, not combat.”* I use this question if a student’s writing conveys a coarse tone of voice. I will use variations of the following question when I observe that a student’s writing is self-focused instead of reader-centered: *“Put yourself in the reader’s position; what would you want to see*,” or “*what would you expect to read here*?” Also, “*how can we better structure this for reader clarity?*” Or, “*how can we better answer readers’ questions*?” These are a few example questions I use to help students expand their thought process and re-examine their writing choices. These questions, along with others, are intended to encourage students to engage in critical inquiry, develop personal agency, foster intellectual self-reliance, and ultimately cultivate self-trust. Once they formulate and propose possible solutions, I then assist them as needed in bringing the solutions into fruition; it is here that I observe students becoming more confident in their abilities and more trusting of their own thoughts, knowledge, ideas, and abilities. From this reflective space, they begin to build self-efficacy, become more motivated and attentive to their writing, and become better thinkers—thinking beyond my feedback—and thoughtful implementers of the feedback they receive. This growth, even though small at times, propels them and gives them a blueprint, so to speak, for revising their writing not only in the present but also in the future.

Post-Feedback DRFS

To gather student feedback about the DRFS, I make a voluntary, anonymous, and ungraded DRFS survey available in a learning management system (LMS). The survey is available to students after the assignment due date and after all final assignment drafts have been submitted in the LMS. Students have the option to provide feedback on the DRFS they attend for all assignments. The post-feedback survey asks students to discuss the feedback they received during one or multiple DRFS, how/if it helped them, and how/if they used the feedback to refine their draft before the final submission. Obtaining feedback from the students helps me further understand their perception and interpretation of the feedback, as well as their ability and comfort level in applying it independently, outside of our DRFS, to enhance their entire paper before submission. I use the survey feedback to gauge their perceptions of the benefits and value of DRFS and to enhance the DRFS.

DRFS Monitoring

At the beginning of each semester, I outline the benefits of DRFS and how they are an essential and complementary component of the teaching and learning process in my classes. I highlight the benefits of DRFS to promote my commitment to student engagement and academic success, to encourage students to take advantage of the opportunity for individualized assistance, and to encourage them to be active participants in their learning. Some students do not take advantage of the DRFS. I am aware of this because I monitor the students who attend the DRFS and their progress from their draft to the final product. Monitoring the students who attend the sessions helps me identify participation, follow-through, engagement with feedback, and application of feedback. Not surprisingly, students who attend DRFS are able to better understand assignment criteria, are more receptive to feedback on their work (because they experience it as assistive rather than solely corrective), and can apply writing concepts, principles, and strategies more effectively. In general, the students who attend the DRFS perform better than those who do not.

Conclusion

Feedback has been a topic researched from various vantage points. The DRFS have been a central element of my teaching. After lectures, discussions, and during-class activities, working one-on-one with students during DRFS has been where my teaching has been most effective and rewarding. During these sessions, I witness students' talents and strengths, as well as their struggles, fear, nervousness, anxiety, and their desire to be seen, heard, respected, and helped. I learn a great deal from students during the DRFS, gaining insights into any modifications I need to make to my discussions, teaching approaches, expectations, and requirements. All of it is done for the benefit of promoting student learning and academic success.

Disclaimer

None.

References

Anderson, P. V. Technical Communication: A Reader-Centered Approach, 9th ed. Boston, MA, USA: Cengage, 2018.

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development* (Vol. 96). Multilingual Matters.

Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment and Evaluation in Higher Education*, *38*(6), 698–712. https://doi.org/10.1080/02602938.2012.691462

Brookhart, S. M. (2008). Feedback that fits. *Educational Leadership*, *65*(4), 54.

Bush, J. C. (2021). Using screencasting to give feedback for academic writing. *Innovation in Language Learning and Teaching*, *15*(5), 473–486. https://doi.org/10.1080/17501229.2020.1840571

Cárcamo, B. (2020). Classifying written corrective feedback for research and educational purposes: A typology proposal. *PROFILE Issues in Teachers' Professional Development*, *22*(2), 211–222. https://doi.org/10.15446/profile.v22n2.79924

Carhill-Poza, A., Tian, L., & Rubio, J. (2023). The affordances of language choice in negotiation of oral peer feedback on L2 writing. *Writing & Pedagogy*, *15*(3), 229–255. https://doi.org/10.1558/wap.21859

Cheng, G. (2022). Exploring the effects of automated tracking of student responses to teacher feedback in draft revision: evidence from an undergraduate EFL writing course. *Interactive Learning Environments*, *30*(2), 353–375. https://doi.org/10.1080/10494820.2019.1655769

Cook, S. (2013). Providing feedback on student writing. https://www.sjsu.edu/aanapisi/docs/ProvidingFeedbackonStudentWritingbySaraCook.pdf

DePaul University. (June 18, 2025). High-Stakes assignments. https://resources.depaul.edu/teaching-commons/teaching-guides/feedback-grading/Pages/high-stakes-assignments.aspx

Ellis, R. (2010). Epilogue: A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, *32*(2), 335–349. https://doi.org/10.1017/S0272263109990544

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, *63*(2), 97–107. https://doi.org/10.1093/elt/ccn023

Fathman, A. K., & Whalley, E. (1990). Teacher response to student writing: focus on form versus content. In B. Kroll (Ed.), *Second Language Writing (Cambridge Applied Linguistics): Research Insights for the Classroom* (pp. 178–190). chapter, Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9781139524551.016 [published online October 5, 2012]

Furneaux, C., Paran, A., & Fairfax, B. (2007). Teacher stance as reflected in feedback on student writing: An empirical study of secondary school teachers in five countries. IRAL - *International Review of Applied Linguistics in Language Teaching, 45*(1), 69–94. https://doi.org/10.1515/IRAL.2007.003

Gombert, S., Fink, A., Giorgashvili, T., Jivet, I., Di Mitri, D., Yau, J., Frey, A., & Drachsler, H. (2024). From the automated assessment of student essay content to highly informative feedback: a case study. *International Journal of Artificial Intelligence in Education*, *34*(4), 1378–1416. https://doi.org/10.1007/s40593-023-00387-6

Han, Y. (2019). Written corrective feedback from an ecological perspective: The interaction between the context and individual learners. *System (Linköping)*, *80*, 288–303. https://doi.org/10.1016/j.system.2018.12.009

Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of Second Language Writing*, *30*, 31–44. https://doi.org/10.1016/j.jslw.2015.08.002

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hopkins, D., & Reid, T. (2024). The academic skills handbook: Your guide to success in writing, thinking and communicating at university.

Hwang, S. (2025). Unpacking the impact of writing feedback perception on self-regulated writing ability: The role of writing self-efficacy and self-regulated learning strategies. *Behavioral Sciences*, *15*(2), 100. https://doi.org/10.3390/bs15020100

Hyland, F. (2003). Focusing on form: Student engagement with teacher feedback. *System*, *31*(2), 217–230. https://doi.org/10.1016/S0346-251X(03)00021-6

Klein, P. D., & Boscolo, P. (2016). Trends in research on writing as a learning activity. *Journal of Writing Research*, *7*(3), 311–350. https://doi.org/10.17239/jowr-2016.07.03.01

Kolin, P. S. (2023). *Successful writing at work* (12th ed.). Cengage.

Lusher, J., Clements, H., & Stevens, E. (2021). A qualitative insight into time-poor/grade-hungry health studies students' perceptions of using assessment criteria and feedback in assignment writing. *Nurse Education Today*, *104*, Article 104999. https://doi.org/10.1016/j.nedt.2021.104999

MacRae, P. (2019). *Business and professional writing: A basic guide* (2nd ed.). Broadview Press.

Mahboob, A. (2015). Understanding and providing 'cohesive' and 'coherent' feedback on writing. *Writing & Pedagogy*, *7*(2–3), 355–376. https://doi.org/10.1558/wap.v7i2-3.26461

Mao, Z., & Lee, I. (2024). Student engagement with written feedback: Critical issues and way forward. *RELC Journal*, *55*(3), 810–818. https://doi.org/10.1177/00336882221150811

Marlink, J., & Wahleithner, J. (2011). *Improving students' academic writing: Building a bridge to success.* Final Report for National Writing Project Local Sites Research Initiative. Retrieved from National Writing Project website: http://www. nwp. org/cs/public/download/nwp_file/15419/ISAW_LSRI_final_report.pdf.

McDonald, D., Boddy, J., O'Callaghan, K., & Chester, P. (2015). Ethical professional writing in social work and human services. *Ethics and Social Welfare*, 9(4), 359–374. https://doi.org/10.1080/17496535.2015.1009481

McLeod, S. H. (2020). Writing across the curriculum: The second stage, and beyond. In *Landmark essays on writing across the curriculum* (pp. 79–86). Routledge.

Moore, J. (2014). *Feedback on academic writing – Part One.* Oxford University Press. https://teachingenglishwithoxford.oup.com/2014/08/28/feedback-on-academic-writing-part-one/

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: perceptions of students and academics. *Assessment and Evaluation in Higher Education*, *42*(2), 266–288. https://doi.org/10.1080/02602938.2015.1103365

Nguyen, L. T. T. (2019). A Case study of teacher feedback on Thai University students' essay writing. *GEMA Online® Journal of Language Studies*, *19*(2), 121–138. https://doi.org/10.17576/gema-2019-1902-08

Nguyen, C.-D. (2024). Scaffolding student engagement with written corrective feedback: Transforming feedback sessions into learning affordances. *Language Teaching Research : LTR*, *28*(5), 1918–1939. https://doi.org/10.1177/13621688211040904

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 (Dorchester-on-Thames), *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Rofiqoh, M. I. (2020). Written and oral feedback in writing class. *RETAIN: Journal of Research in English Language Teaching*, *8*(2).

Tayebipour, F. (2019). The impact of written vs. oral corrective feedback on omani part-time vs. full-time college students' accurate use and retention of the passive voice. *Journal of Language Teaching and Research*, *10*(1), 150. https://doi.org/10.17507/jltr.1001.17

Tehrani, F. A. (2018). Feedback for writing or writing for feedback? *The Journal of Language and Linguistic Studies*, *14*(4), 162.

Williams, J. G., & Jasmine, G. (2003). Providing feedback on ESL students' written assignments. *The Internet TESL Journal*, *9*(10), 1–5.

Zhang, Z. (Victor), & Hyland, K. (2022). Fostering student engagement with feedback: An integrated approach. *Assessing Writing*, *51*, Article 100586. https://doi.org/10.1016/j.asw.2021.100586

Chapter 7

The Cognitive Load Theory, Element Interactivity and Mathematics Education

Bing Hiong Ngu*
and Huy P. Phan
University of New England, Armidale, NSW, Australia

Abstract

Element interactivity is a fundamental concept within the framework of cognitive load theory. It refers to the interaction between elements that must be processed simultaneously in order for meaningful understanding to occur. Anything that requires learning—such as a number, symbol, a concept or procedure—is considered an element. Element interactivity is a measure of task complexity, which is influenced by individual differences. The level of element interactivity reflects not only the inherent complexity of the task but also the learner's level of expertise. Experts can process multiple interactive elements as a single unit, thereby reducing the perceived complexity of the task. Consequently, a task that is complex for a novice may be relatively simple for an expert. While cognitive load researchers typically use element interactivity to assess the intrinsic complexity of learning materials, we argue that this concept is equally applicable when evaluating the complexity of *instructional materials*. Previous research has examined instructional materials designed as variability tasks (e.g., word problems sharing a similar problem structure but differ in contexts). High variability tasks are characterized by high element interactivity, requiring learners to invest greater germane cognitive load, which in turn enhances learning

* Corresponding Author's Email: bngu@une.edu.au

In: Progress in Education. Volume 91
Editor: Roberta V. Nata
ISBN: 979-8-90134-375-3

outcomes. Importantly, such instructional materials are particularly beneficial for learners with higher levels of prior knowledge.

In this book chapter, we review several of our empirical studies in mathematics education, which demonstrate the value of incorporating element interactivity as a lens for evaluating not only the complexity of learning materials, but also the relative effectiveness of different instructional approaches across various secondary school mathematics topics. Our findings suggest that instructional approaches with low element interactivity—and hence lower cognitive load—are often more effective particularly for learning complex learning materials. Furthermore, we highlight the critical role of leveraging prior knowledge to reduce the complexity of instructional materials.

Introduction

In the foundational work of Sweller (1994) and Sweller and Chandler (1994), the concept of element interactivity within cognitive load theory (Sweller et al., 2011; Sweller et al., 2019) was introduced primarily to define the complexity of learning materials. Element interactivity refers to the extent to which individual elements within learning material must be processed simultaneously to achieve understanding and learning. An element is defined as any unit of information that requires learning (e.g., a number, symbol, procedure, etc.). Paas and Van Merriënboer (1994) demonstrated that high-variability tasks—such as geometrical problems that differ not only values but also problem formats—are associated with high element interactivity and require learners to invest greater germane cognitive resources, which in turn enhances learning outcomes. In a recent review, Chen et al. (2023) identified element interactivity as a source of cognitive load for comparing the relative effectiveness of worked examples and problem solving. They argue that higher levels of element interactivity associated with problem solving—due to the need to search for a solution path—can impose extraneous cognitive load, which is unproductive for learning. In contrast, studying worked examples enables learners to focus on the solution steps, which involve low element interactivity and thereby facilitate learning.

In this chapter, we review a number of our prior empirical studies that highlight how element interactivity can serve not only as a measure of material complexity but also as a criterion for evaluating the relative effectiveness of instructional approaches in mathematics education. In contrast to earlier research that identifies high element interactivity in problem-solving tasks as a source of extraneous cognitive load (Chen et al., 2023) and associates high-variability tasks with germane cognitive load (Paas & Van Merriënboer, 1994), we investigate element interactivity as a basis for evaluating instructional effectiveness from the perspective of intrinsic cognitive load. Specifically, when comparing two instructional approaches for teaching a particular mathematics topic—each differing in element interactivity—we argue in favor of designs with lower element interactivity. Such approaches reduce the cognitive demands on learners and support more effective learning. Since element interactivity is a fundamental construct in cognitive load theory, we will discuss cognitive load theory and human cognitive architecture in the next section.

Cognitive Load Theory and Human Cognitive Architecture

Cognitive Load Theory distinguishes between two types of knowledge that humans can acquire (Geary, 2007, 2008; Geary, 2012): biologically primary knowledge and biologically secondary knowledge. Biologically primary knowledge is acquired naturally through immersion in a specific cultural context—for example, learning a native dialect. In contrast, biologically secondary knowledge is acquired primarily through explicit instruction in formal educational settings, such as learning how to solve mathematics problems.

According to Cognitive Load Theory, there are five principles that align with the architecture of human cognition and support the acquisition of biologically secondary knowledge (Sweller et al., 2011; Sweller et al., 2019).

1. Information Store Principle
   This principle highlights the role of long-term memory as a vast repository for domain-specific knowledge stored in the form of schemas (Tricot & Sweller, 2014). Schemas that hold in the long-term memory determine an individual's level of expertise. Therefore, a primary goal of instruction is to facilitate the acquisition of schemas.
2. Borrowing and Reorganizing Principle
   This principle emphasizes that a highly efficient method of acquiring biologically secondary knowledge is by borrowing information from others (e.g., teachers, texts, experts).
3. Randomness as Genesis Principle

This principle refers to the process of randomly generating and testing problem-solving strategies in the absence of relevant schemas.

4. Narrow Limits of Change Principle
This principle underscores the limitations of working memory, both in terms of capacity (Cowan, 2001) and duration (Peterson & Peterson, 1959), when processing new information. Instructional design should take these constraints into account to prevent cognitive overload.
5. Environmental Organizing and Linking Principle
This principle explains how environmental cues can trigger the retrieval of relevant schemas from long-term memory, enabling individuals to generate appropriate responses.

Element Interactivity and Three Types of Cognitive Loads

According to Sweller (2010), there are three types of cognitive load that influence instructional design. The first is intrinsic cognitive load, which is determined by the level of *element interactivity* relative to a learner's expertise in a given domain. As noted above, element interactivity refers to the extent to which elements within learning materials must be processed simultaneously to enable learning. The more complex the material, the higher the element interactivity—and therefore the greater the intrinsic load. However, as learners gain expertise, the perceived element interactivity decreases, reducing the intrinsic cognitive load (Kalyuga et al., 1998; Sweller, 2024).

The second type is extraneous cognitive load, which results from poorly designed instructional materials. This type of load is unnecessary and interferes with learning by increasing element interactivity in ways that do not support schema acquisition. For instance, asking learners to solve problems without guidance—such as through means-ends analysis to search for a solution path —can impose extraneous cognitive load and hinder schema development (Ngu & Phan, 2023; Sweller & Cooper, 1985).

The third type is germane cognitive load, which refers to the cognitive effort devoted to processes that are essential for learning. It is associated with element interactivity that supports the construction of schemas, making it an intrinsic part of effective learning. For example, practicing with varied problem types increases element interactivity and, consequently, germane cognitive load; however, it benefits learning (Likourezos et al., 2019).

As noted earlier, the inherent complexity of learning material is governed by intrinsic cognitive load, which, in turn, is determined by the level of element interactivity. We argue that instructional material should be regarded as a subset of learning material, as learners must first understand its presentation and then apply it to solving specific types of problems. The next section will examine the relationship between the complexity of both the learning materials and the instructional materials, as well as element interactivity, across three mathematics topics.

Element Interactivity and Learning Materials

The complexity of learning material can be estimated by counting the number of elements and the interactions among them (Chen et al., 2023; Sweller, 1994; Sweller & Chandler, 1994). In earlier work, Sweller (1994) estimated the element interactivity involved in solving an algebraic problem such as *a/b = c, solve for a* as comprising five elements (i.e., *a*, /, *b*, =, and *c*). However, Chen et al. (2023) estimated a total of 28 elements for the same algebraic problem. Unlike Sweller's (1994) approach, Chen et al. (2023) included, among other ways of counting, the elements explicitly present in the problem statement itself. For example, the expression *a/b = c, solve for a* was considered to consist of six distinct elements: *a, /, b, =, c*, plus *solve for a.*

Leahy and Sweller (2008) conducted an experiment on teaching elementary students how to read school bus timetables, in which the step "Look up the route numbers" was counted as a distinct element. In another study, Leahy and Sweller (2019) estimated element interactivity in the context of teaching students how to write puzzle poems. For example, the 1st element was "To write one, you need to follow all these rules," followed by the 2nd element, "There must be six lines in the poem," and so on.

In our previous research in mathematics education, specifically within the domain of linear equations (Ngu et al., 2015; Ngu & Phan, 2016; Ngu & Phan, 2021; Ngu & Phan, 2022; Ngu et al., 2023; Ngu et al., 2018), trigonometry problems (Ngu & Phan, 2023), and percentage problems (Ngu et al., 2016; Ngu et al., 2023; Ngu et al., 2018; Ngu et al., 2014), we identified two types of elements: (i) numerical and symbolic elements (e.g., numbers, symbols such as /, %, $), and (ii) mathematical concepts. To estimate the level of element interactivity, we used two criteria: the presence of individual elements and the relationships among them, which we refer to as mathematical concepts.

Linear Equations

In our prior research, we distinguished between the relational line and the operational line when describing the solution procedure for linear equations (e.g., Ngu & Phan, 2022). The relational line represents the quantitative relationship between the left and right sides of an equation. In contrast, the operational line depicts the application of a mathematical operation (e.g., +3 on both sides) that transforms the state of the equation while preserving its equality. Drawing on the topic of linear equations (Ngu & Phan, 2021), we use the balance method to illustrate different levels of element interactivity by comparing a one-step equation (Table 1) with a two-step equation (Table 2). The two-step equation involves a higher level of element interactivity and is therefore more complex.

Table 1. Element count for a one-step linear Equation

One-step linear equation

Line 1	$5x = 10$	(÷5 on both sides)
Line 2	÷5 ÷5	
Line 3	$x = 2$	

Solution step		Element	Concept	Total element
Line 1	$5x = 10$	$5x$, 10	• x is a pronumeral • $5x$ means 5 multiplied by x • The '=' sign describes the quantitative relation between elements whereby the left side of the equation equals its right side • To solve for x, the learner performs the same operation on both sides of the equation to maintain its equality	6
Line 2	÷5 ÷5	÷5	• Cancel ÷5 with ×5 in the left side • The same ÷5 must interact with 10 on the right side to maintain its equality	3
Line 3	$x = 2$	x, 2	• 2 is the solution	3
				12

Note: a concept is regarded as an element.

As shown in Table 1, the one-step equation consists of two relational lines (Lines 1 and 3) and one operational line (Line 2). Line 1 contains six elements: two mathematical elements ($5x$, 10) and four conceptual elements (e.g., recognizing x as a pronumeral). Line 2 includes three elements: one mathematical operation (÷5) and two conceptual elements (e.g., Cancel ÷5 with ×5 in the left side). Line 3 consists of three elements: two mathematical elements (x, 2) and one conceptual element (i.e., recognizing 2 as the solution). In total, the one-step equation comprises 12 elements.

Table 2. Element count for a two-step linear equation

Two-step linear equation

Line 1	$3x + 1 = 10$	(–1 on both sides)
Line 2	–1 –1	
Line 3	$3x = 9$	(÷3 on both sides)
Line 4	÷3 ÷3	
Line 5	$x = 4$	

Solution step		Element	Concept	Total element
Line 1	$3x + 1 = 10$	$3x$, 1, 10	• x is a pronumeral • $3x$ means 3 multiplied by x • The '=' sign describes the quantitative relation between elements whereby the left side of the equation equals its right side • To solve for x, the learner performs the same operation on both sides of the equation to maintain its equality	7
Line 2	–1 –1	–1	• Cancel –1 with + 1 in the left side • The same –1 must interact with 10 on the right side to maintain its equality	3

Note: Lines 3, 4, and 5 resemble the structure of a one-step equation, collectively contributing 12 elements. Consequently, the total number of elements for the two-step equation (i.e., $3x + 1 = 10$) amounts to $12 + 10 = 22$.

As presented in Table 2, the two-step equation comprises three relational lines (Lines 1, 3, and 5) and two operational lines (Lines 2 and 4). Notably, Lines 3, 4, and 5 in the two-step equation are similar to Lines 1, 2, and 3 in the one-step equation, and therefore share the same number of elements—12 in total. As such, the total number of elements in the two-step equation: number of elements in Lines 1 and 2 plus 12 elements in Lines 3, 4, and 5. Using the same method of counting, Line 1 includes seven elements: three mathematical elements ($3x$, 1, 10) and four conceptual elements (e.g., recognizing x as a pronumeral). Line 2 contains three elements: one mathematical operation (–1) and two conceptual elements (e.g., Cancel –1 with + 1 in the left side). Therefore, the total number of elements in the two-step equation: 7 + 3 + 12 = 22 elements.

In summary, the two-step equation contains more relational lines (3 vs. 2) and operational lines (2 vs. 1) than the one-step equation. A subset of the two-step equations (Lines 3, 4, and 5) shares a similar problem structure—and consequently, similar solution steps—with the one-step equation (Lines 1, 2, and 3). As confirmed by the element count, the one-step equation involves fewer elements than the two-step equation (12 vs. 22). Therefore, the two-step equation demonstrates a higher level of element interactivity compared to the one-step equation.

Trigonometry Problems

Trigonometry problems resemble linear equations involving fractions (e.g., $\cos 56° = x/18$ vs. $30 = x/5$). These problems can be classified into two types based on the position of the pronumeral (numerator vs. denominator), which influences their complexity—for example, $\tan 35° = x/20$ (Type I) versus $\sin 42° = 70/x$ (Type II). Trigonometry problems and linear equations with fractions share a similar problem structure. As shown in Tables 1 and 2, linear equations with more solution steps tend to have a higher number of relational and operational lines, which increases their element count. Consequently, instead of counting assumed elements and their interactions, we estimated the element interactivity of trigonometry problems using the number of relational and operational lines. A higher number of relational and operational lines indicates greater element interactivity and, therefore, increased complexity. As indicated in Table 3, a Type I problem—where the pronumeral is in the numerator—has fewer relational lines (2 vs. 3) and operational lines (1 vs. 2) than a Type II problem—where the pronumeral is in the denominator. Thus, Type II problems are more complex due to their higher level of element interactivity.

Table 3. Element count for two types of trigonometry problems

Type I problem		Type II problem	
Line 1	$\tan 35° = \frac{x}{20}$ [÷ 20 becomes × 20]	Line 1	$\sin 42° = \frac{70}{x}$ [$\div x$ becomes $\times x$]
Line 2	$x = \tan 35° \times 20$	Line 2	$x \times \sin 42° = 70$
Line 3	$x = 14.00$	Line 3	$x \sin 42° = 70$ [$\times \sin 42°$ becomes $\div \sin 42°$]
		Line 4	$x = \frac{70}{\sin 42°}$
		Line 5	$x = 104.61$

Note:
Type I problem: The pronumeral (unknown) is in the numerator. It involves two relational lines (Lines 1 and 3) and one operational line (Line 2).
Type II problem: The pronumeral (unknown) is in the denominator. It involves three relational lines (Lines 1, 3, and 5) and two operational lines (Lines 2 and 4)

Percentage Problems

Apart from estimating the relative levels of element interactivity for linear equations and trigonometry problems, we have also examined different levels of element interactivity for both simple percentage problems (Ngu et al., 2023) and complex percentage problems (Ngu et al., 2018). To illustrate how we estimate the level of element interactivity for a simple percentage problem, we analyzed the number of elements involved in the solution process for a simple percentage question used in Ngu et al.'s (2023) study: "*A worker pays $350 in tax per month, which is 8% of her monthly income. What is her monthly income?*" (see Table 4).

Table 4. A simple percentage problem

Question: A worker pays $350 in tax per month, which is 8% of her monthly income. What is her monthly income?

Solution Procedure		
Step 1	Create an unknown	Let x be the monthly income
Step 2	8% of monthly income is $350	$8\% \times x = \$350$
Step 3	Solve the equation	$x = \$350 \div 8$
		$x = \$4{,}375$

Solution step		Element	Concept	Total element
Step 1	Let x be the monthly income	x	• x is a variable	2
Step 2	$8\% \times x = \$350$	8%, x, $350	• The relationship among percentage, quantity, and the variable, such that the left side of the equation equals the right side.	4
Step 3	$x = \$350 \div 8\%$ $= \$4{,}375$	x, $350, 8%	• Understand multiplication (×) as the inverse operation of division (÷) in order to solve for x	4
				10

Using the algebra approach, the solution procedure has three steps.

Step 1: assign a variable (x) to represent the unknown monthly income.

Step 2: integrate the variable (x) with the given values (8% and $350) to construct an equation such as $8\% \times x = \$350$.

Step 3: solve for x.

Table 5. A complex percentage problem

Question: After a 12% markup, the shoes now cost \$34. How much did they originally cost?

Solution Procedure		
Step 1	Create an unknown	Let x be the original cost
Step 2	Markup price = original cost + increased amount	$\$34 = x + 12\% \times x$
Step 3	Factorize	$\$34 = x(1 + 12\%)$
Step 4	Solve the equation	$x = \$34 \div (1 + 12\%)$
		$x = \$30.35$

Solution step		Element	Concept	Total element
Step 1	Let x be the original cost	x	• x is a variable	2
Step 2	$\$34 = x + (12\% \times x)$	\$34, x, 12%	• the product of $12\% \times x$ represents an increased amount • the sum of $x + (12\% \times x)$ represents the marked up price of \$34	5
Step 3	$\$34 = x(1 + 12\%)$	\$34, x, (1 +12%)	• the expression (1 +12%) represents 100% plus 12% • The product x (1 + 12%) is the result of factorizing $x + 12\% x$ in step 2 • The product x (1 + 12%) is equal to \$34	6
Step 4	$x = \$34 \div (1 + 12\%)$	x, \$34, (1 +12%)	• Divide \$34 by (1 + 12%) to obtain the value of x, which is the solution	4
				17

Step 1: Assign a variable (x) to represent the original cost.

Step 2: Integrate the variable (x) with the values (\$34 and 12%) and express the relationship in an equation: $\$34 = x + (12\% \times x)$.

Step 3: Factorize the expression "$x + 12\%x$" to produce the equation:
$\$34 = x \times (1 + 12\%)$.

Step 4: Solve for x.

As shown in Table 4, the number of elements involved in steps 2, 2, and 3 are 1, 4, and 4, respectively, resulting in a total of 10 elements.

In our prior research (Ngu et al., 2018), we also estimated the level of element interactivity for a complex percentage problem (see Table 5): "*After a 12% markup, the shoes now cost $34. How much did they originally cost?*" (Parker & Leinhardt, 1995, p 4). Once again, we used the algebra approach to generate the solution procedure, which consists of four steps:

As shown in Table 5, the number of elements corresponding to Steps 1, 2, 3, and 4 are 2, 5, 6, and 4, respectively, resulting in a total of 17 elements.

In summary, an examination of the solution procedures for the percentage problems (simple vs. complex) reveals that a component of the complex problem—specifically, the expression (12% × x)—shares a similar structural pattern with the simple problem (8% × x). The complex problem can be described as a 'double-reference problem,' in which the unknown variable (x) appears twice in the equation—a feature that poses a challenge for secondary school students (Koedinger et al., 2008). As revealed by the element count, the complex problem involves more elements than the simple one (17 vs. 10), making it inherently more complex. Having examined how element interactivity can be used to distinguish the complexity of learning materials, the next section will explore its role in evaluating the relative effectiveness of instructional approaches.

Element Interactivity and Instructional Design

As noted previously, this chapter highlights how element interactivity can be used to assess both the complexity of learning materials and the effectiveness of instructional designs. From the perspective of instructional design and cognitive load, previous research has framed element interactivity as contributing to either extraneous cognitive load (Chen et al., 2023) or germane cognitive load (Paas & van Merrienboer, 1994). For instance, worked examples reduce extraneous cognitive load by providing solution steps, while problem-solving increases it through unguided search for a solution. Conversely, task variability promotes germane cognitive load by encouraging deeper processing of essential components of the material, ultimately enhancing learning outcomes.

In a study by Lu et al. (2020), element interactivity was used to distinguish three instructional formats for teaching adult students how to write Chinese characters. In the isolated-integrated format, students traced and practiced isolated components before progressing to the full character; the variability-integrated format varied the positions of the components; and the integrated-integrated format involved tracing and practicing full characters throughout. Tracing and practicing full characters from the start imposed higher element interactivity. Both the isolated-integrated and variability-integrated groups outperformed the integrated-integrated group, with the variability group performing best. The authors attributed this to reduced element interactivity from the isolated elements effect, enabling learners to manage increased intrinsic load introduced by component variation—consistent with the variability of worked examples effect (Paas & van Merrienboer, 1994). However, we argue that both the isolated-integrated and variability-integrated groups likely experienced similar levels of element interactivity, with the variability-integrated group benefiting from learning the spatial positioning of components, a critical factor in mastering full Chinese character writing.

Aside from the three studies outlined above, no other prior research has examined the use of element interactivity to assess the relative effectiveness of instructional approaches. In the next section, we present our previous studies that explore this issue from an element interactivity perspective—that is, by treating element interactivity as a source of intrinsic cognitive load.

Linear Equations

In the teaching of linear equations, the balance method is commonly used in Western countries, whereas the inverse method is the preferred method in Asian countries (Ngu et al., 2015; Ngu & Phan, 2016; Ngu et al., 2023). The main difference between the balance method and inverse method lies in the operational line (e.g., – 6 on both sides, a balance operation vs. + 6 becomes – 6, an inverse operation) (Table 6). With respect to the operational line, the interaction between elements occurs on both sides of the linear equation in the balance method, but only on one side of the linear equation in the inverse method. Therefore, a balance operation incurs twice as many interactive elements compared to an inverse operation for each operational line; consequently, the balance method imposes a higher level of element interactivity than the inverse method. However, the advantage of the inverse method over the balance method depends on the complexity of linear equations (Ngu et al., 2015; Ngu et al., 2018; Ngu et al., 2022). The inverse method has no advantage over the balance method for one-step equations which involve one operational line and two relational lines, thereby imposing low cognitive load (e.g., Ngu et al., 2015). However, the inverse method is better than the balance method as the complexity of linear equations increases (Ngu et al., 2018). The inverse method is particularly effective as compared to the balance method for linear equations that involve multiple solution steps such as linear equations with negative pronumeral and negative number (e.g., $7 - 2x = 11$) (Ngu et al., 2022).

Table 6. The balance method and inverse method

Balance method		Inverse method	
Line 1	$x + 6 = 8$ (-6) on both sides	Line 1	$x + 6 = 8$ ($+6$ becomes -6)
Line 2	$-6 \quad -6$	Line 2	$x = 8 - 6$
Line 3	$x = 2$	Line 3	$x = 2$

Prior research has highlighted the influence of prior knowledge on instructional effectiveness (Ngu et al., 2023; Ngu et al., 2018). From the perspective of element interactivity, prior knowledge can reduce the number of interacting elements within a familiar method, thereby lowering cognitive load and enhancing learning. In a cross-cultural study by Ngu et al. (2023), Australian and Malaysian students were randomly assigned to either the balance or inverse method on learning linear equations. In Australia, the balance group outperformed those in the inverse group, likely due to their prior exposure to the balance method. Conversely, Malaysian students in the inverse group outperformed their Australian peers, reflecting their prior experience with the inverse method.

Trigonometry Problems

In a previous study, we compared the problem-solving approach, worked examples approach, and analogy approach for learning to solve trigonometry problems from the perspective of element interactivity (Ngu & Phan, 2023). Consistent with Chen et al. (2023), we argued that varying levels of element interactivity favor the worked examples approach. Searching for a solution path through problem solving imposes a higher level of element interactivity than studying worked examples, which present step-by-step solutions.

The design of the analogy approach is grounded in the theory of learning by analogy, which involves leveraging prior knowledge of a learned problem to facilitate understanding of a new problem with a similar problem structure and solution procedure. Together, the learned and new problems form a pair of isomorphic problems. Successful one-to-one mapping of elements and their relationships between a pair of isomorphic problems enables learners to infer a shared problem structure and, consequently, a shared solution procedure.

As shown in Table 7, a pair of isomorphic problems comprise a linear equation with a fraction and a trigonometry problem. Both problems contain the same number of elements and share two relational lines and one operational line. Learners can establish a one-to-one mapping between the relational elements (e.g., mapping *m* to *x*, tan 75° to 2, and so on), which is expected to support the identification of a shared solution procedure. From an element interactivity perspective, concurrently processing such a pair of isomorphic problems (i.e., two worked examples) is likely to impose approximately twice the level of element interactivity—and thus cognitive load—compared to studying worked examples sequentially, as in the worked examples approach. However, prior knowledge of linear equations with fractions may allow students to process isomorphic worked examples with lower element interactivity, thereby reducing cognitive load. Specifically, students may treat a linear equation with a fraction as a single element or chunk, thereby lowering the overall level of element interactivity in the analogy approach. This rationale was supported in our findings, which contradicted our initial prediction that the analogy approach would impose a higher level of element interactivity compared to the worked examples approach (Ngu & Phan, 2023).

Table 7. A pair of isomorphic examples

A linear equation with a fraction	A trigonometry problem with pronumeral as a numerator
$\frac{m}{7} = 2$ [÷7 becomes ×7]	$\tan 75° = \frac{x}{26}$ [÷ 26 becomes × 26]
$m = 2 \times 7$	$x = \tan 75° \times 26$
$m = 14$	$x = 97.03$

Note: A sample of material used in Ngu and Phan (2023).

Percentage Problems

Prior research indicates distinct cultural preferences in problem-solving strategies. For example, Chinese students often use the algebra approach (e.g., setting up equations to solve for unknowns), while U.S. students tend to favor the concrete method (e.g., using diagrams) (Cai, 2000). Similar contrasts appear in percentage problem-solving. For example, in solving percentage problems such as "*A tank contained 120 L of water, which was 5% of its total capacity. What is the total capacity*?", Australian textbooks (e.g., Vincent et al., 2012) emphasize the unitary method where the unit percentage is the key concept — calculating 1% and then multiplying by 100. In contrast, Singaporean textbooks (e.g., Chow, 2007; Toh et al., 2019), reflecting Chinese instructional practices, emphasize the algebra approach (e.g., 5% × x = 120 L, solve for x).

We previously compared the algebra and unitary approaches for solving percentage problems such as: "*A worker pays $350 in tax per month, which is 8% of her monthly income. What is her monthly income?*" (Ngu et al., 2023). Using the concept of element interactivity as a framework, we evaluated the relative effectiveness of the algebra versus non-algebra (Unitary) approaches in supporting students' acquisition of percentage problem-solving skills.

Table 8. Unitary approach

Question: A worker pays $350 in tax per month, which is 8% of her monthly income. What is her monthly income?

Solution Procedure		
Step 1		8% of monthly income = $350
Step 2	1% of monthly income	$350 ÷ 8 = $43.75
Step 3	100% of monthly income	$43.75 × 100 = $4,375

Solution step		Element	Concept	Total element
Step 1	8% of monthly income = $350	8%, monthly income, $350,	• the relationship between 8%, monthly income, and $350 where the left side of the equation equals to its right side	4
Step 2	$350 ÷ 8 = $43.75	$350, 8, $4.375	• $350 represents 8% of monthly income • the sub-goal is 1% of monthly income	5
Step 3	$43.75 × 100 = $4,375	$43.75, 100, $4375	• $43.75 is equivalent to 1% of monthly income • $43.75 × 100 is equivalent to 100% of monthly income	5
				14

As shown in Table 4 and Table 8, the element count for the equation approach is 10, which is fewer than the unitary approach, 14. Therefore, different levels of element interactivity favor the equation approach. Specifically, in our study, we examined the relative effectiveness of the algebra approach and unitary approach across two levels of learner expertise (low prior knowledge vs. high prior knowledge) (Ngu et al., 2023). The findings indicate no significant difference between the equation and unitary approaches for either simple or complex problems, regardless of students' prior knowledge. This suggests that, at each level of prior knowledge, the difference in element interactivity between the two approaches may not be large enough to influence performance outcomes. In contrast, high prior knowledge students outperformed their low prior knowledge peers on both simple and complex problems. These findings suggest that high prior knowledge students are able to process both the equation and unitary approaches with fewer interacting elements. This is supported by the observation that low prior knowledge students reported greater mental effort across both approaches.

Discussion

When Sweller (1994) first introduced the concept of element interactivity, it was used solely to describe the complexity of learning materials. It enables researchers to quantify complexity based on the number of elements and the interactions among them. Importantly, it also highlights the need to define material complexity relative to the learner's level of expertise (Sweller, 1994; Sweller & Chandler, 1994)—an aspect largely overlooked in prior research (Chen et al., 2023).

In this chapter, drawing on our prior research across three mathematical topics—linear equations, trigonometry, and percentage problems—we have highlighted the distinction between simple and complex problems based on element count and, consequently, element interactivity. Our key contribution lies in clearly identifying the number and types of elements involved, as well as the interactions among them. To estimate element interactivity, we applied two criteria: the presence of individual elements and the relationships among them (i.e., mathematical concepts). For each topic, we distinguish between simple and complex problems in terms of element interactivity, with complex problems involving a greater number of elements and, therefore, higher element interactivity than simple problems.

Prior research has identified element interactivity associated with problem-solving as a source of extraneous cognitive load (Chen et al., 2023), whereas element interactivity arising from variability tasks is considered a source of germane cognitive load (Paas & Van Merriënboer, 1994; Lu et al., 2020; Likourezos et al., 2019). In contrast, we emphasized the use of element interactivity as a framework for evaluating the relative effectiveness of different instructional approaches from the perspective of intrinsic cognitive load—an area that has received limited attention in cognitive load research. We reviewed the design of various instructional materials across linear equations, trigonometry, and percentage problems, where the same mathematical concepts were presented differently (e.g., unitary approach vs. equation approach in the domain of percentage problems), resulting in varying levels of element interactivity. In the context of linear equations, for example, both the balance method and the inverse operation address the same concept—applying a mathematical operation to change the equation while preserving equality. However, the inverse operation is more effective because it imposes a lower level of element interactivity. In particular, the inverse method is more effective than the balance method for multi-step linear equations (e.g., Ngu & Phan, 2022).

In the domain of trigonometry, the analogy approach demonstrates how leveraging prior knowledge—such as understanding a linear equation involving fractions—can support learning a structurally similar trigonometric problem, thereby facilitating transfer through a shared solution strategy (Ngu & Phan, 2023). In comparing the algebra and unitary approaches to solving percentage problems, our findings highlight the advantage of the algebra approach, which involves fewer interacting elements and is therefore more cognitively efficient (Ngu et al., 2023). Although the current *Australian Curriculum: Mathematics* advocates the unitary method—a non-algebra approach—our research, grounded in cognitive load theory, supports the use of algebra instruction to enhance students' learning of percentage problems.

In this chapter, we reviewed the application of element interactivity across three mathematics topics within the context of algebra instruction—namely, linear equations, trigonometry problems, and percentage problems. Specifically, we employed the concept of element interactivity to evaluate not only the complexity of the learning materials but also the instructional effectiveness of different teaching approaches for these topics. A natural question arises as to whether this approach can be extended beyond algebra to other areas of mathematics. For instance, in the domain of geometry, a trapezium can be understood as comprising two component shapes: a triangle and a parallelogram. There are two common ways to calculate its area:

1. $A = \frac{1}{2}(a + b) \times h$ where a and b are the lengths of the two parallel sides, and h is the height (i.e., the perpendicular distance between them).
2. By decomposing the trapezium into a triangle and a parallelogram, $A = \frac{1}{2} \times b_1 h + b_2 h$ where b_1 is the base of the triangle, and b_2 is the base of a parallelogram, and h is the shared height.

This example illustrates how element interactivity might also be applied to geometry, offering a pathway to explore the cognitive demands of instructional approaches across a broader range of mathematical domains.

In most mathematics curricula, students are typically taught how to calculate the area of a triangle and a parallelogram separately before learning how to calculate the area of a trapezium. Therefore, it is pedagogically sound to introduce the trapezium area formula by leveraging students' prior knowledge of these two shapes, thereby reducing element interactivity involved in learning the trapezium area formula. An important instructional question then arises: once students have mastered the second method—calculating the area of a trapezium by decomposing it into a triangle and a parallelogram—should instruction proceed to the first method? The benefit of the first method is that it offers a single, more streamlined formula for calculating the area, potentially simplifying future problem-solving. Future research could empirically compare these two instructional methods based on element interactivity, examining which method more effectively supports learning to calculate the area of a trapezium, and whether introducing the first method after the second provides additional benefits.

Conclusion

This chapter highlights the theoretical value of element interactivity in quantifying task complexity and cognitive load. Practically, it provides a powerful lens for evaluating instructional effectiveness in mathematics education. Instructional designs with lower element interactivity better support novice learners, while leveraging prior knowledge can reduce element interactivity. Therefore, this book chapter provides valuable insights into the design of mathematics curriculum. Tailoring instruction based on cognitive load theory and aligning it with human cognitive architecture enhances learning outcomes across diverse mathematical topics.

References

Chen, O., Paas, F., & Sweller, J. (2023). A Cognitive load theory approach to defining and measuring task complexity through element interactivity. *Educational Psychology Review*, *35*(2), 63. https://doi.org/10.1007/s10648-023-09782-w

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci*, *24*(1), 87-114. https://doi.org/10.1017/S0140525X01003922

Geary, D. C. (2007). Educating the evolved mind: Conceptual foundations for an evolutionary educational psychology. In J. S. Carlson & J. R. Levin (Eds.), *Psychological perspectives on contemporary educational issues* (pp. 1-99). Information Age Publishing.

Geary, D. C. (2008). An Evolutionarily Informed Education Science. *Educational Psychologist*, *43*(4), 179-195. https://doi.org/10.1080/00461520802392133

Geary, D. C. (2012). Evolutionary Educational Psychology. In K. Harris, S. Graham, & T. U (Eds.), *APA Educational Psychology handbook* (Vol. 1, pp. 597–621). American Psychological Association.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors*, *40*(1), 1-17. https://doi.org/10.1518/001872098779480587

Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade-Offs between grounded and abstract representations: Evidence From algebra problem solving. *Cognitive Science*, *32*(2), 366-397. https://doi.org/10.1080/03640210701863933

Leahy, W., & Sweller, J. (2008). The imagination effect increases with an increased intrinsic cognitive load. *Applied Cognitive Psychology*, *22*(2), 273-283. https://doi.org/10.1002/acp.1373

Leahy, W., & Sweller, J. (2019). Cognitive Load Theory, Resource Depletion and the Delayed Testing Effect. *Educational Psychology Review*, *31*(2), 457-478. https://doi.org/10.1007/s10648-019-09476-2

Ngu, B. H., Chung, S. F., & Yeung, A. S. (2015). Cognitive load in algebra: element interactivity in solving equations. *Educational Psychology*, *35*(3), 271-293. https://doi.org/10.1080/01443410.2013.878019

Ngu, B. H., & Phan, H. P. (2016). Unpacking the complexity of linear equations from a cognitive load theory perspective. *Educational Psychology Review*, *28*, 95-118. https://doi.org/10.1007/s10648-015-9298-2

Ngu, B. H., & Phan, H. P. (2021). Learning linear equations: capitalizing on cognitive load theory and learning by analogy. *International Journal of Mathematical Education in Science and Technology*, 1-17. https://doi.org/10.1080/0020739X.2021.1902007

Ngu, B. H., & Phan, H. P. (2022). Advancing the study of solving linear equations with negative pronumerals: A smarter way from a cognitive load perspective. *PLOS One*, *17*(3), e0265547.

Ngu, B. H., & Phan, H. P. (2023). Differential instructional effectiveness: overcoming the challenge of learning to solve trigonometry problems that involved algebraic transformation skills. *European journal of psychology of education*, *38*(4), 1505-1525. https://doi.org/10.1007/s10212-022-00670-5

Ngu, B. H., Phan, H. P., Hong, K. S., & Usop, H. (2016). Reducing intrinsic cognitive load in percentage change problems: The equation approach. *Learning and Individual Differences*, *51*, 81-90. https://doi.org/https://doi.org/10.1016/j.lindif.2016.08.029

Ngu, B. H., Phan, H. P., Hong, K. S., & Usop, H. (2023). Is the inverse method more effective than the balance method on learning linear equations? A cross-cultural experimental study between Australia and Malaysia. *Asian Journal for Mathematics Education*. https://doi.org/10.1177/27527263231194470

Ngu, B. H., Phan, H. P., Usop, H., & Hong, K. S. (2023). Instructional efficiency: The role of prior knowledge and cognitive load. *Applied Cognitive Psychology*, *37*(6), 1223-1237. https://doi.org/10.1002/acp.4117

Ngu, B. H., Phan, H. P., Yeung, A. S., & Chung, S. F. (2018). Managing Element Interactivity in Equation Solving. *Educational Psychology Review*, *30*(1), 255-272. https://doi.org/10.1007/s10648-016-9397-8

Ngu, B. H., Yeung, A. S., Phan, H. P., Hong, K. S., & Usop, H. (2018). Learning to solve challenging percentage-change problems: A cross-cultural study from a cognitive load perspective. *The Journal of Experimental Education*, *86*(3), 362-385. https://doi.org/10.1080/00220973.2017.1347774

Ngu, B. H., Yeung, A. S., & Tobias, S. (2014). Cognitive load in percentage change problems: unitary, pictorial, and equation approaches to instruction. *Instructional Science*, *42*(5), 685-713. https://doi.org/10.1007/s11251-014-9309-6

Paas, F., & Van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, *86*(1), 122-133. https://doi.org/http://dx.doi.org/10.1037/0022-0663.86.1.122

Parker, M., & Leinhardt, G. (1995). Percent: A privileged proportion. *Review of Educational Research*, *65*(4), 421-481. https://doi.org/10.3102/00346543065004421

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*(3), 193-198. https://doi.org/http://dx.doi.org/10.1037/h0049234

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295-312. https://doi.org/http://dx.doi.org/10.1016/0959-4752(94)90003-5

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*(2), 123-138. https://doi.org/10.1007/s10648-010-9128-5

Sweller, J. (2024). Cognitive load theory and individual differences. *Learning and Individual Differences*, *110*, 102423. https://doi.org/10.1016/j.lindif.2024.102423

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer. doi:10.1007/978-1-4419-8126-4.

Sweller, J., & Chandler, P. (1994). Why some material Is difficult to learn. *Cognition and Instruction*, *12*(3), 185-233. https://doi.org/10.1207/s1532690xci1203_1

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*(1), 59-89. https://doi.org/10.1207/s1532690xci0201_3

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*(2), 261-292. https://doi.org/10.1007/s10648-019-09465-5

Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review*, *26*(2), 265-283. https://doi.org/10.1007/s10648-013-9243-1

Chapter 8

Peer Mentorship and a Novice ESL Teacher's Professional Development

Norhaiza Hamdan[1], Med
Juliana Othman[2], PhD
and Lo Yueh Yea[2,*], PhD
[1]Language Department, IPG Kampus Perempuan Melayu, Melaka, Malaysia
[2]Department of Language and Literacy Education,
University of Malaya, Malaysia

Abstract

This chapter focuses on the unique challenges and opportunities that novice ESL teachers encounter as they begin their educational careers. It provides a holistic framework for professional growth, emphasizing the development of essential teaching skills, classroom management strategies, and reflective practices. It highlights the critical role of mentorship, collaborative learning communities, and ongoing professional training in fostering confidence and competence in new educators. By integrating research-based insights with practical tools, this chapter equips novice ESL teachers with the knowledge and strategies necessary to navigate the complexities of the second language classroom, foster meaningful relationships with students, and adapt to the evolving demands of the education sector. Ultimately, this chapter aims to empower early-career teachers to cultivate a sustainable and fulfilling professional journey, laying the foundation for long-term success and impact in the field of ESL education.

---

* Corresponding Author's Email: norhaiza@ipgm.edu.my

In: Progress in Education. Volume 91
Editor: Roberta V. Nata
ISBN: 979-8-90134-375-3

Keywords: ESL novice teachers, professional development, peer mentoring, reflective practice, teaching skills

Introduction

Peer mentoring is defined as a professional development strategy that bridges the gap between teacher preparation and the remainder of an educator's career, facilitating enhanced instructional practices, shifts in professional mindset, and collaborative interactions within professional learning communities. In the ESL context, peer mentoring offers an alternative approach to observational learning, enabling novice teachers to acquire effective instructional strategies from colleagues while receiving targeted feedback on their professional development. The critical need for targeted professional development support stems from the recognition that ESL teachers face unique challenges, including language barriers, cultural differences, and varying proficiency levels among students, making continuous professional development essential to ensure educators remain effective, adaptable, and motivated.

The chapter starts by introducing the theoretical foundations on which this study was based. It then addresses the challenges faced by novice ESL teachers before focusing on the benefits of the peer mentoring approach as they transition from training to the workplace. In doing so, the chapter outlines the implementation strategies, the challenges, and the future direction of this professional development approach.

Theoretical Foundation

The efficacy of peer mentoring programs is grounded in situated learning theory, which recognizes that learning is situational and can occur within the context of real-life experiences. Peers who have experienced similar training and have been working in the school context are more capable of giving richer meanings to novice teachers during the mentoring partnership. These novice teachers benefit most from experiences that are context-based and collaborative, rather than didactic, with training and assistance strategies that have been proven most effective. Communities of Practice (CoP), a model developed by Lave and Wenger (1991) within the framework of Situated

Learning Theory, offers another crucial theoretical lens. This model emphasizes that peer collaboration enhances teaching practices and improves student outcomes, with professional learning communities fostering a culture of shared knowledge and support. The role of collaborative professional development in teacher retention cannot be overstated. Professional Learning Communities offer teachers a convenient way to share best practices and brainstorm innovative approaches to enhance learning and drive student achievement. Peer mentoring offers possibilities for more reciprocal roles, where novice teachers and their mentors become reciprocal co-learners. These theoretical foundations converge to suggest that peer mentoring represents a powerful mechanism for supporting novice ESL teachers through the critical transition from preparation to practice.

The Landscape of Novice ESL Teachers' Challenges

The transition from pre-service training to real-world ESL classrooms presents novice teachers with a complex array of professional, institutional, and personal challenges that significantly impact their early career development and retention. Recent research indicates that ESL teaching positions are among the most difficult to fill in U.S. schools, with 59% of elementary schools and 69% of high schools reporting difficulties filling ESL or bilingual education vacancies for the 2024-25 academic year (National Center for Education Statistics, 2024). This shortage reflects not only the high demand for qualified ESL educators but also the significant challenges that drive teachers away from the profession during their crucial early years.

Linguistic Complexity

Novice ESL teachers face the formidable challenge of managing multiple language proficiency levels within single classrooms, requiring sophisticated pedagogical differentiation skills that often exceed their initial preparation. Francis (2019) emphasizes that effective ESL instruction requires teachers to create environments where students "embrace their culture and their language as a foundation of who they are" while simultaneously developing English proficiency. This linguistic complexity manifests in diverse ways: teachers must simultaneously address beginning-level students who may have limited literacy in any language, intermediate learners developing academic English, and advanced students refining sophisticated language skills for academic success. The challenge is intensified by the reality that many ESL students have limited opportunities to use English outside the classroom, leading to reduced motivation and engagement (Nguyen et al., 2023). Teachers must therefore create authentic, meaningful contexts for language use while managing the cognitive load of instruction across multiple proficiency levels.

Cultural Responsiveness

Navigating diverse cultural backgrounds and learning styles represents another significant challenge for novice ESL teachers. Recent research emphasizes the critical importance of culturally competent relationships, where teachers invest in understanding the particulars of students' lives and incorporate these elements into lessons (Toppel, cited in Edutopia, 2019). However, many novice teachers lack the experiential knowledge to bridge cultural differences effectively and may struggle to create inclusive environments that honor students' diverse backgrounds. The challenge extends beyond surface-level cultural awareness to a more profound understanding of how cultural values impact learning preferences, family engagement expectations, and student behavioral norms. Novice teachers often report feeling unprepared to address personal issues among students with varying educational backgrounds and the complex socio-emotional needs of students navigating cultural transitions.

Assessment Challenges

Distinguishing between language acquisition stages and academic ability represents one of the most complex professional demands facing novice ESL teachers. This challenge requires a sophisticated understanding of second language acquisition theory combined with practical assessment literacy. Teachers must differentiate between students who lack content knowledge and those who need additional language support to demonstrate their understanding. The complexity is compounded by the need to provide meaningful feedback that supports both language development and academic growth, while also preparing students for high-stakes assessments that may not accurately reflect their capabilities during the phases of language acquisition.

Curriculum Adaptation

Modifying mainstream content for English language learners requires a deep understanding of content combined with language teaching expertise. Gonzalez (cited in Edutopia, 2019) emphasizes that all teachers must understand their role as language teachers and become aware of the specialized academic language within their content areas. However, novice ESL teachers often struggle to balance content delivery with language instruction, particularly when working with a curriculum that assumes native-level English proficiency.

Limited Preparation

A significant gap exists between teacher preparation programs and classroom realities, with most novice teachers experiencing sudden discontinuation of support after initial certification (Farrell, 2012). Recent research with pre-service EFL teachers revealed that while university coursework provided fundamental knowledge of theories and methodologies, it failed to prepare them adequately for classroom realities, offering only "an idealistic view of the classroom" (Çelik, 2019). This theory-practice disconnect leaves novice teachers struggling to apply academic knowledge in complex, dynamic teaching environments. The challenge is particularly acute in ESL contexts where teachers need specialized knowledge about second language acquisition, cultural responsiveness, and differentiated instruction that may not be adequately addressed in general teacher preparation programs.

Professional Isolation

Often, being the sole ESL specialist in schools creates profound professional isolation for novice teachers. This isolation limits opportunities for peer collaboration, instructional feedback, and professional growth that are critical during the early career phase. Research in the Vietnamese context demonstrates that novice EFL teachers struggle particularly with classroom management and student engagement when lacking peer support and mentorship (Tran et al., 2024).

Benefits of Peer Mentoring for Novice ESL Teachers

The implementation of peer mentoring programs for novice ESL teachers yields substantial benefits across multiple dimensions of professional development, emotional well-being, institutional integration, and ultimately, student learning outcomes. Recent systematic reviews have demonstrated that peer mentoring in educational contexts produces favorable impacts across various outcomes, including academic performance, retention rates, emotional and psychological well-being, and social integration (Le et al., 2024). These benefits are particularly pronounced in ESL contexts where teachers face unique challenges related to linguistic diversity, cultural responsiveness, and complex pedagogical demands.

Pedagogical Content Knowledge

Peer mentoring facilitates the development of ESL-specific teaching strategies through collaborative observation and shared practice. Hamdan et al. (2024) demonstrate that novice ESL teachers develop professionally through observing peers during mentoring relationships, with a particular emphasis on role modeling that shapes both their attitudes and instructional approaches. This observational learning enables novice teachers to witness effective ESL-specific strategies in action, including scaffolding techniques for multilingual learners, differentiated instruction for varying proficiency levels, and culturally responsive teaching practices. The collaborative nature of peer mentoring enables real-time feedback and the adaptation of teaching strategies. Professional learning communities and peer mentoring programs facilitate the exchange of best practices, foster experimentation, and reduce the isolation that often accompanies individual attempts at innovation (Novik et al., 2024). Through structured teacher networks, novice ESL educators gain access to a repository of proven instructional strategies while contributing their own fresh perspectives and innovative approaches.

Classroom Management

Managing multilingual, multicultural classrooms requires a sophisticated understanding of diverse learning styles, cultural norms, and communication patterns. Peer mentoring offers novice ESL teachers practical strategies for creating inclusive environments that value linguistic diversity while maintaining effective learning structures. Recent research emphasizes the importance of mindful grouping strategies that provide English language learners with opportunities to collaborate with both peers who share their native language and native English speakers (Colorín Colorado, 2021). Experienced peer mentors share culturally responsive classroom management techniques that address the specific challenges of multilingual environments, including strategies for managing code-switching, facilitating peer learning across language barriers, and creating safe spaces for linguistic risk-taking. These approaches acknowledge that students may have specific cultural norms around peer interaction and require a nuanced understanding of diverse communication styles.

Assessment Literacy

Peer mentoring enhances novice teachers' understanding of formative and summative assessment in ESL contexts, particularly the complex task of distinguishing between language acquisition stages and academic ability. Through collaborative reflection and shared expertise, mentoring relationships help novice teachers develop assessment literacy that accounts for linguistic development while accurately measuring content knowledge. Experienced peers share practical strategies for implementing alternative assessment approaches, including portfolio-based assessment, performance assessments, and multilingual evaluation techniques, which provide more accurate representations of student learning. This collaborative approach to assessment development ensures that novice teachers develop both theoretical understanding and practical skills in ESL-appropriate assessment practices.

Technology Integration

The rapid evolution of digital technologies has emphasized the need for enhanced teacher competencies in implementing interactive learning tools.

Collaborative teacher networks play a significant role in advancing digital pedagogy, with peer mentoring programs facilitating knowledge sharing about effective technology integration (Tan, 2024). These networks encourage teachers to reflect on their practices, co-develop lesson plans incorporating digital tools, and collectively troubleshoot technical challenges. Peer mentoring provides structured opportunities for sharing digital tools and resources specifically designed for language learning, including adaptive learning platforms, virtual reality immersion programs, and AI-powered language learning applications. Recent innovations demonstrate that integrating technology through peer collaboration can lead to significant improvements in student engagement and language acquisition outcomes.

Stress Reduction

Peer mentoring serves as an efficient and cost-effective mechanism for helping novice teachers cope with challenges, reduce stress, and prevent burnout during their critical early career phase (Singh et al., 2014; Pölczman et al., 2024). The shared problem-solving approach inherent in peer mentoring relationships provides emotional support while developing practical solutions to common ESL teaching challenges. Research demonstrates that mentoring relationships provide both mentors and mentees with enhanced social support networks, contributing to improved psychological well-being and reduced feelings of professional isolation. The collaborative nature of peer mentoring creates opportunities for shared reflection on stressful situations and collective development of coping strategies.

Professional Confidence

Peer mentoring has a significant impact on teacher self-efficacy through multiple pathways, including vicarious experiences, verbal persuasion, and emotional regulation support. Burger (2024) demonstrates that constructivist mentoring approaches have positive effects on teacher self-efficacy, with enhanced self-awareness and problem-solving abilities contributing to increased professional confidence. The validation and encouragement provided through peer relationships help novice teachers develop realistic expectations while building confidence in their abilities. Mentors' verbal persuasions play an important role in self-efficacy development, particularly when combined with collaborative partnership approaches that emphasize mutual learning rather than hierarchical knowledge transfer (Kutsyuruba et al., 2019).

Networking

Peer mentoring facilitates the development of professional relationships and career connections that extend beyond formal mentoring periods. These networks provide ongoing support for professional growth and create opportunities for collaborative innovation in ESL education. The 47% of companies implementing employee mentoring and coaching programs demonstrates the widespread recognition of peer learning as a practical approach to professional development (LinkedIn, 2024).

Professional networks developed through peer mentoring relationships often evolve into communities of practice that continue supporting teachers throughout their careers, providing resources for advanced professional development and leadership opportunities.

Work-Life Balance

Mentoring relationships provide practical strategies for managing the demanding nature of ESL teaching while maintaining personal well-being. Peer mentors share time management techniques, stress reduction strategies, and boundary-setting approaches that help novice teachers develop sustainable professional practices. The emotional support provided through peer relationships helps teachers navigate the psychological demands of working with diverse student populations while maintaining their own mental health and professional effectiveness.

Implementing Effective Peer Mentoring Programs

The successful implementation of peer mentoring programs for novice ESL teachers requires comprehensive planning, systematic design, and ongoing support structures that address both the technical and relational aspects of mentoring relationships. Recent research emphasizes that effective mentoring programs must be carefully designed to address individual requirements and challenges of participants, with successful implementation dependent on multiple interconnected factors, including participant matching, program structure, and technological support (Gehreke et al., 2024).

Needs Assessment

Effective peer mentoring programs begin with a thorough needs assessment that identifies specific challenges and goals for novice ESL teachers within their particular institutional contexts. The National Mentoring Resource Center (2022) emphasizes that program coordinators must conduct comprehensive evaluations of participant needs before designing mentoring interventions. This assessment should examine multiple dimensions, including instructional challenges, emotional support needs, institutional integration requirements, and professional development goals. For ESL teachers specifically, a needs assessment must address unique challenges, such as managing multilingual classrooms, developing cultural responsiveness, navigating complex assessment requirements, and adapting the curriculum for diverse learners. Hamdan et al. (2024) demonstrate that novice ESL teachers benefit particularly from observational learning opportunities, where they can witness effective teaching strategies in action. This suggests that needs assessment should identify specific areas where classroom observation and modeling would be most beneficial. The assessment process should involve multiple stakeholders, including novice teachers, experienced educators, administrators, and student representatives, to ensure a comprehensive understanding of program requirements. Data collection methods may include surveys, focus groups, individual interviews, and analysis of existing performance data to create detailed profiles of participant needs and institutional contexts.

Participant Selection

The careful selection of both mentors and mentees is a critical factor in program success, with research indicating that successful matching between mentors and mentees significantly influences program effectiveness (Fladerer et al., 2023). For mentor selection, programs should establish clear criteria that include not only teaching experience and expertise but also communication skills, cultural competency, and commitment to collaborative learning relationships. Effective mentor selection criteria should encompass pedagogical expertise in ESL instruction, demonstrate success in multicultural classroom management, strong interpersonal communication skills, and enthusiasm for professional learning and growth. Research suggests that mentors with shared values and similar improvement orientations create more effective mentoring relationships, particularly in terms of professional learning and growth outcomes. Mentee selection should focus on identifying novice teachers who demonstrate openness to learning, willingness to engage in reflective practice, and commitment to professional growth. Programs should also consider factors such as teaching assignments, grade levels, student populations, and geographic proximity when feasible to ensure relevant and practical mentoring relationships.

Program Structure

The choice between formal and informal approaches, along with decisions about duration and intensity, significantly impacts program effectiveness. Gehreke et al. (2024) found that both individual and group-based peer mentoring formats demonstrate positive outcomes. However, programs should be designed to address the specific needs of participants rather than adopting one-size-fits-all approaches. Formal structures provide clear frameworks with defined roles, expectations, and accountability measures, while informal approaches offer flexibility and organic relationship development. Many successful programs combine elements of both, establishing formal frameworks for participant matching, goal setting, and progress monitoring while allowing flexibility in meeting formats, communication frequency, and activity selection. Duration considerations should strike a balance between sufficient time for relationship development and practical scheduling constraints. Research suggests that mentoring relationships require a minimum period of 3-6 months to develop trust and achieve meaningful outcomes, with optimal programs extending for full academic years to support comprehensive professional development. Intensity decisions involve frequency of interactions, time commitments, and scope of activities. Effective programs typically include weekly or bi-weekly structured interactions supplemented by informal communication and collaboration opportunities.

Mentor Training

Practical mentor training develops both coaching and communication skills essential for successful peer relationships. Guethler (2024) emphasizes that mentoring programs should incorporate goal-setting frameworks and strategic planning workshops to help mentors develop proactive approaches to relationship building and professional development support. Training components should include the development of communication skills, active listening techniques, effective feedback delivery methods, conflict resolution strategies, and cultural competency development. ESL-specific training should address second language acquisition theory, culturally responsive teaching practices, assessment strategies for multilingual learners, and techniques for supporting teachers working with diverse student populations. The training should also cover mentoring relationship dynamics, boundary setting, confidentiality protocols, and strategies for balancing support with autonomy. Practical training elements should include role-playing exercises, case study analysis, and opportunities to practice mentoring techniques in supervised settings.

Orientation for Participants

Comprehensive orientation programs establish clear expectations and ground rules for all participants. Practical orientations address program goals, participant roles and responsibilities, communication protocols, confidentiality requirements, and available resources and support systems. Orientation should include relationship development strategies, goal-setting processes, conflict resolution procedures, and evaluation methods. Participants need a clear understanding of program timelines, meeting expectations, documentation requirements, and processes for addressing challenges or concerns.

Goal Setting Processes

Structured goal-setting processes create focus and accountability for both individual participants and program-level objectives. Individual goal setting should involve collaborative development of specific, measurable, achievable, relevant, and time-bound (SMART) objectives that align with participant needs assessment findings and professional development priorities. Program-level objectives should address both participant outcomes (knowledge development, skill acquisition, confidence building) and institutional goals (retention improvement, teaching effectiveness enhancement, culture development). Regular review and adjustment of goals ensures continued relevance and motivation throughout the mentoring relationship.

Technology-Enhanced Mentoring

Digital mentoring platforms have revolutionized accessibility and effectiveness of mentoring relationships, with 85% of mentors and mentees reporting that digital tools significantly enhance the mentoring experience through increased flexibility and communication opportunities (Deloitte, 2024). Video conferencing platforms enable face-to-face interaction regardless of geographic constraints, while collaborative tools facilitate real-time document sharing and project development.

Platforms like Zoom, Google Meet, and MS Teams provide comprehensive communication tools, progress tracking, and administrative dashboards. These platforms offer customizable templates, automated scheduling, integrated calendar systems, and robust reporting capabilities that streamline program management while enhancing participant experience.

Digital libraries and teaching material repositories enable efficient sharing of resources, lesson plans, assessment tools, and professional development materials. Cloud-based platforms, such as Google Workspace or Microsoft Office 365, provide collaborative document editing capabilities that support real-time feedback and joint lesson development. Resource sharing platforms should include organized libraries of ESL-specific materials, assessment tools, professional development resources, and links to external learning opportunities. Version control and access management ensure the appropriate sharing of information while protecting intellectual property and maintaining security.

Professional learning communities and forums extend mentoring relationships beyond formal pairings to create broader networks of support and collaboration. Platforms like LinkedIn Learning, industry-specific forums, and institutional learning management systems provide opportunities for ongoing professional development and peer connection. These networks enable participants to access diverse expertise, participate in group discussions, share challenges and successes, and maintain connections beyond formal program periods. Integration with formal mentoring programs enhances both individual relationships and broader professional community building.

Mobile apps provide convenient access to mentoring resources, communication tools, and progress tracking capabilities. Apps enable quick communication, calendar integration, resource access, and notification systems that support ongoing engagement without requiring dedicated computer time. Effective mobile integration includes push notifications for meeting reminders, quick messaging capabilities, mobile-optimized resource libraries, and offline access to key materials. User-friendly interfaces and reliable functionality ensure high adoption rates and consistent engagement throughout mentoring relationships. The successful implementation of peer mentoring programs requires comprehensive attention to design, training, implementation, and the integration of technology. With careful planning and ongoing support, these programs offer effective mechanisms for supporting novice ESL teachers while fostering sustainable professional learning communities.

Challenges and Solutions in Peer Mentoring Programs

The implementation and maintenance of effective peer mentoring programs, while highly beneficial, present numerous challenges that require strategic solutions and continuous adaptation. Research demonstrates that mentoring relationships can face significant obstacles, including communication barriers, time management issues, and mismatched expectations, with mentors particularly struggling to balance their responsibilities with mentoring duties (Pölczman et al., 2024). Understanding these challenges and implementing evidence-based solutions is crucial for program sustainability and effectiveness.

Balancing mentoring responsibilities with existing teaching demands represents one of the most persistent challenges in peer mentoring programs. Recent research suggests that time management issues often hinder mentoring relationships, as mentors struggle to fulfill their academic responsibilities while providing sufficient support to mentees (Cho & Lee, 2021). The challenge is particularly acute for ESL teachers who already face complex demands, including curriculum adaptation, multilingual classroom management, and extensive family communication requirements. Time constraints manifest in multiple ways, including difficulty scheduling regular meetings, inadequate time for preparation and reflection, competing priorities that interrupt mentoring activities, and burnout resulting from overcommitment. These challenges often lead to inconsistent engagement and reduced program effectiveness, ultimately compromising the quality of support provided to novice teachers.

Managing mismatched mentor-mentee relationships poses significant challenges for both program administrators and participants. Research emphasizes that personality factors play important roles in mentoring relationship success, with characteristics such as shyness and limited communication inversely correlated with positive outcomes (LaFleur & White, 2010). Successful matching requires consideration of multiple factors, including communication styles, professional goals, cultural backgrounds, and personal preferences. Personality conflicts can manifest as communication breakdowns, differing expectations about relationship dynamics, incompatible working styles, and fundamental disagreements about teaching approaches or professional values. These mismatches can lead to participant dissatisfaction and early termination of the relationship, reducing program effectiveness and potentially creating negative experiences that discourage future participation.

Maintaining engagement throughout program duration poses ongoing challenges for both individual relationships and overall program sustainability. Research indicates that willingness to participate in mentoring programs can vary significantly, with digital formats showing notably lower engagement rates compared to face-to-face interactions (Davishahl et al., 2021). Inconsistent participation undermines relationship development and reduces the cumulative benefits of sustained mentoring support. Factors contributing to inconsistent participation include changing priorities and commitments, lack of perceived value or benefit, inadequate program structure and support, and competing demands on participants' time and energy. These challenges require proactive strategies to maintain motivation and engagement throughout program cycles.

Working within budget and staffing constraints significantly impacts program design and implementation capacity. Many programs struggle with limited funding for mentor training, insufficient technological infrastructure, inadequate administrative support, and a lack of dedicated coordination resources. These limitations require that programs rely heavily on volunteer participation and minimal support structures, which may compromise program quality and sustainability.

Practical Solutions and Best Practices

Creative approaches to finding collaboration time have emerged as essential solutions for addressing time constraints. Successful programs implement multiple scheduling options, including virtual meetings, micro-mentoring sessions, asynchronous communication platforms, and flexible meeting formats that accommodate varying schedules and preferences. Modern mentoring platforms offer automated scheduling tools that seamlessly integrate with existing calendar systems, thereby reducing administrative burdens while maximizing scheduling efficiency. Some programs implement "mentoring on demand" models that allow participants to connect for brief consultations as needed, supplementing formal scheduled meetings with flexible support options.

Allowing mentees to work with different experts addresses both personality conflicts and diverse learning needs. Research supports the implementation of mentoring circles, team mentoring, and network approaches, which provide multiple sources of support and expertise (Chronus, 2024). This approach reduces pressure on individual relationships while providing broader perspectives and resources. Programs can implement mentor rotation systems, subject-specific expert panels, or peer learning networks that enable mentees to access diverse expertise while reducing dependency on single mentoring relationships.

Building mentoring into job descriptions and evaluation systems creates institutional support and accountability structures that enhance program sustainability. This integration includes formal recognition in professional development requirements, performance evaluation criteria, and workload allocation systems. Successful integration requires leadership commitment to mentoring as a professional responsibility, clear expectations for participation and performance, and systematic support systems that enable effective participation without creating additional burden.

Regular program evaluation and refinement processes enable responsive adaptation to emerging challenges and changing needs. Effective evaluation systems combine quantitative metrics (such as participation rates, retention statistics, and goal achievement) with qualitative feedback (including relationship quality, participant satisfaction, and learning outcomes) to provide a comprehensive assessment of the program.

Continuous improvement requires systematic data collection, regular feedback loops, responsive program adjustments, and a commitment to evidence-based decision-making. Programs implementing these approaches demonstrate higher participant satisfaction and improved outcomes over time. These challenges and solutions highlight the complexity of implementing effective peer mentoring programs while providing practical guidance for addressing common obstacles. Success requires sustained commitment to both participants and program quality, supported by evidence-based practices and responsive adaptation to emerging needs.

Future Directions and Recommendations

The landscape of peer mentoring for novice ESL teachers is evolving rapidly, driven by technological innovations, changing educational paradigms, and emerging research insights. As the field advances toward 2030 and beyond, several key trends and recommendations will shape the future of mentoring programs, requiring strategic adaptation and forward-thinking implementation approaches.

AI-supported mentoring platforms represent a transformative frontier in the implementation of peer mentoring. The global AI in education market is projected to grow from $5.18 billion in 2024 to $112.3 billion by 2034, with AI mentoring platforms offering unprecedented opportunities for personalized learning experiences (World Economic Forum, 2025). Recent developments demonstrate AI's capacity to enhance mentor-mentee matching through sophisticated algorithms that consider skills, interests, goals, and even emotional intelligence factors, delivering bias-free, accurate connections that optimize relationship outcomes (Qooper, 2024). AI mentoring software provides 24/7 access to coaching support, virtual mentoring assistants, and personalized resource recommendations, while automated administrative functions streamline program management and reduce the coordinator's workload (MentoringComplete, 2024). However, research emphasizes that AI should complement rather than replace human mentors, focusing on augmenting relationship quality through data-driven insights and administrative efficiency rather than substituting the essential human elements of empathy and contextual understanding.

Brief, targeted support interactions are emerging as valuable supplements to traditional mentoring relationships, providing just-in-time guidance that addresses immediate needs without requiring extensive time commitments. This approach leverages digital platforms to enable quick consultations, targeted skill development sessions, and on-demand access to experts, accommodating busy professional schedules while maintaining meaningful support connections.

International teacher exchange and collaboration programs are expanding through digital platforms that transcend geographical limitations, enabling ESL teachers to access diverse expertise and cultural perspectives. These global mentoring networks offer opportunities for cross-cultural learning that enhances cultural responsiveness and broadens professional understanding of multilingual education approaches across diverse international contexts.

Including student feedback in mentoring conversations represents an emerging trend that enhances teaching effectiveness through direct learner input. Research demonstrates strong connections between responsiveness to student voice and academic outcomes, with students providing valuable insights into instructional effectiveness that can inform mentor-mentee discussions and professional development goals (Kahne et al., 2022).

Policy and Practice Recommendations

Integrating peer mentoring concepts into pre-service education requires systematic curriculum revision that includes mentoring theory, relationship-building skills, and collaborative learning approaches. Teacher preparation programs should incorporate field-based experiences that expose candidates to peer mentoring models, develop communication and feedback skills essential for future participation, and establish foundations for lifelong professional learning through collaborative relationships. Administrative policies must facilitate and sustain mentoring programs through dedicated funding, time allocation, and institutional integration. Successful district support includes building mentoring expectations into job descriptions and evaluation systems, providing released time for mentoring activities, and creating recognition systems that value collaborative professional development contributions. Including mentoring competencies in teacher evaluation systems establishes institutional expectations for collaborative professional learning while providing clear standards for mentoring skill development. These standards should address communication skills, feedback delivery, cultural competency, and knowledge-sharing capabilities that enhance both individual and collective professional capacity. Sustainable program financing requires diversified funding strategies that combine federal, state, local, and private resources. Successful approaches include demonstrating a return on investment through retention improvements, leveraging teacher quality grants, and establishing partnerships with higher education institutions and professional organizations that share the costs and benefits of mentoring programs.

Conclusion

Recent research provides compelling evidence of the transformative power of peer mentoring in supporting the professional development of novice ESL teachers. Hamdan et al. (2024) revealed that peer mentoring plays a pivotal role in the professional growth of novice ESL teachers, particularly through mechanisms of observational and collaborative learning, demonstrating how this approach facilitates enhanced instructional practices, shifts in professional mindset, and collaborative interactions within professional learning communities.

The evidence base supporting the effectiveness of peer mentoring continues to strengthen. A comprehensive meta-analysis found that induction and mentoring programs had statistically significant effects on teacher and student outcomes, including retention, efficacy, and achievement. In contrast, research examining variation in mentoring aspects found that time spent with a mentor, participation in mentor-facilitated professional development activities, and the quality of mentors' interactions with novice teachers were related to novice teachers' perceptions of mentoring success, self-reflection, and efficacy. Critical factors for successful implementation include structured mentor training, clearly defined roles and responsibilities, regular monitoring of program implementation, and the establishment of trusting relationships between mentors and mentees.

Educational leaders and policymakers must recognize peer mentoring as an essential investment in teacher quality and student success. Successful programs must have a clear purpose and vision for improving teaching and learning, as well as intentional investments in built-in time, training, and the use of an evidence-based instructional rubric. They should also have defined roles, responsibilities, and compensation for mentors. Implementation steps should include conducting comprehensive needs assessments, establishing mentor selection and training protocols, creating structured observation and feedback cycles, and developing sustainability plans that integrate mentoring into broader school improvement initiatives.

For teacher educators, the implications extend beyond individual program design to systemic reform. Programs to support teachers' professional growth and ultimately improve the quality of English language instruction must emphasize ongoing, practical, and collaborative professional development that addresses the unique challenges facing ESL educators in increasingly diverse educational contexts.

The transformative potential of peer mentoring extends far beyond individual teacher development to address broader educational equity and student success goals. Research has shown that peer mentoring programs can significantly improve graduation rates and address equity gaps, particularly for underrepresented student populations. The benefits of peer mentoring can be categorized into four fundamental aspects: academic performance, retention rates, emotional and psychological well-being, and social integration.

As the ESL field continues to evolve, peer mentoring represents a powerful mechanism for building sustainable professional learning communities that not only support individual teacher growth but also contribute to systemic improvements in educational quality and equity. The evidence suggests that when implemented with fidelity and sustained commitment, peer mentoring programs become catalysts for transforming both teaching practice and student outcomes, creating ripple effects that extend throughout educational systems and communities.

References

Burger, C. (2024). The role of constructivist mentoring in the professional development of novice teachers. *Teaching and Teacher Education*, 147, 104819. https://doi.org/10.1016/j.tate.2024.104819

Çelik, S. (2019). Managing the 'unmanageable': The journey of a novice teacher. *Australian Journal of Teacher Education*, 44(4), 16-30. https://doi.org/10.14221/ajte.2018v44n4.2

Cho, Y., & Lee, E. (2021). The impact of time management on mentoring relationships among teachers. *Mentoring & Tutoring: Partnership in Learning*, 29(5), 589–607. https://doi.org/10.1080/13611267.2021.1987223

Chronus. (2024). *Mentoring circles: A guide to implementing group mentoring*. Chronus Corporation.

Colorin Colorado. (2021). *Cooperative learning strategies.* WETA Public Broadcasting. https://www.colorincolorado.org/article/cooperative-learning-strategies

Davishahl, J., Erickson, M., & McPherson, S. (2021). Engagement in digital versus face-to-face mentoring. *Journal of Digital Learning in Teacher Education*, 37(3), 172-185. https://doi.org/10.1080/21532974.2021.1914455

Farrell, T. S. C. (2012). Novice-service language teacher development: Bridging the gap between preservice and in-service education and development. *TESOL Quarterly*, 46(3), 435-449. https://doi.org/10.1002/tesq.36

Fladerer, M. P., Braun, S., & Fasbender, U. (2023). The role of mentor-mentee similarity in mentoring relationships. *Journal of Vocational Behavior*, 145, 103902. https://doi.org/10.1016/j.jvb.2023.103902

Francis, D. (2019). *Culturally responsive teaching in the ESL classroom.* TESOL Press.

Gehreke, J., Miller, A., & Thompson, K. (2024). Designing effective peer mentoring programs: A meta-analysis of structural factors. *Educational Research Review*, 42, 100588. https://doi.org/10.1016/j.edurev.2024.100588

Guethler, P. (2024). *Strategic planning for mentor success.* Association for Talent Development.

Hamdan, N., Othman, J., & Lo, Y. Y. (2024). Enhancing ESL Novice Teachers' Professional Development: Roles of Peers in Observational Learning Through Peer Mentoring. *LLT Journal: A Journal on Language and Language Teaching*, *27*(2), 866-877. https://doi.org/10.24071/llt.v27i2.8411

Kahne, J., Lee, N. J., & Feezell, J. T. (2022). The civic and political significance of student voice. *Teachers College Record*, 124(1), 1-42. https://doi.org/10.1177/01614681221087401

Kutsyuruba, B., Walker, K., & Gooden, A. (2019). The role of mentoring in fostering self-efficacy. *Mentoring & Tutoring: Partnership in Learning*, 27(3), 254-275. https://doi.org/10.1080/13611267.2019.1630990

LaFleur, J., & White, B. J. (2010). Apprehension and the mentoring relationship. *Communication Research Reports*, 27(3), 205-214. https://doi.org/10.1080/08824091003737916

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Le, H., Tran, T., & Nguyen, H. (2024). The impacts of peer mentoring in education: A systematic review. *Review of Educational Research*, 94(1), 112-145. https://doi.org/10.3102/00346543231218864

LinkedIn. (2024). *2024 Workplace learning report*. LinkedIn Learning.

Mentoring Complete. (2024). *AI-powered mentoring software.* Guider Inc.

National Center for Education Statistics. (2024). *Teacher vacancies in U.S. public schools: 2024-25 academic year.* U.S. Department of Education.

National Mentoring Resource Center. (2022). *Elements of effective practice for mentoring. MENTOR: The National Mentoring Partnership*.

Nguyen, T., Brown, J., & Garcia, L. (2023). Motivation and engagement in multilingual classrooms. *Language Teaching Research*, 27(4), 789-812. https://doi.org/10.1177/13621688221123456

Novik, N., Johnson, S. M., & Donaldson, M. L. (2024). Professional learning communities and teacher innovation. *Educational Policy*, 38(1), 45-71. https://doi.org/10.1177/089590482311778

Pölczman, T., Szabó, M., & Kovács, K. (2024). Mentoring as a protective factor against teacher burnout. *Teaching and Teacher Education*, 141, 104490. https://doi.org/10.1016/j.tate.2024.104490

Qooper. (2024). *AI in mentoring: The future of mentorship programs*. Qooper Mentoring Software.

Singh, K., Milne, J., & Walther, M. (2014). The efficacy of mentoring as a stress reduction intervention. *Australian Journal of Guidance and Counseling,* 24(2), 146-161. https://doi.org/10.1017/jgc.2014.14

Tan, E. (2024). Digital pedagogy and collaborative teacher networks. *Computers & Education*, 215, 105033. https://doi.org/10.1016/j.compedu.2024.105033

Tran, L., Pham, H., & Le, V. (2024). Challenges for novice EFL teachers in Vietnam: The role of peer support. *Asia-Pacific Journal of Teacher Education*, 52(1), 88-105. https://doi.org/10.1080/1359866X.2023.2287452

World Economic Forum. (2025). *The future of AI in education: Global market report*. WEF.

Index

Volume 91

Progress in Education

Roberta V. Nata
Editor

nova science publishers
www.novapublishers.com

ISBN 979-8-90134-391-3
9 798901 343913