OPEN PRAXIS

From One-Size Texts to Tailored Readings: Student Experiences with AI-Generated Course Materials

RESEARCH ARTICLES

ALEXANDER M. SIDORKIN

INTERNATIONAL COUNCIL FOR OPEN AND DISTANCE EDUCATION

ABSTRACT

Commercial textbooks impose structural constraints: high cost, fixed reading level, and limited local relevance. Generative AI enables instructors to produce weekly readings tailored to student interests and comprehension levels, but introduces risks around sourcing and unsupported claims. This mixed-methods classroom study documents a proof of principle for AI-generated readings in a graduate educational leadership course, examining practical acceptability and instructional boundaries. The intervention used an AI assistant to generate weekly readings through iterative student prompting, implementing dual tailoring (differentiation) along two dimensions: interest (examples, sector focus, professional role) and comprehension level (pacing, scaffolding, vocabulary density). Data include end-of-course survey responses from 24 students and systematic artifact analysis of weekly reading session logs. Students rated the readings as effective for preparation, with 75 percent agreeing they learned more than in a comparable course without AI assistance. Artifact analysis confirms substantial tailoring through sector anchoring, role framing, and comprehension adjustments triggered by clarification requests. However, readings often lacked internal traceability, citation quality was inconsistent, and over-specific institutional claims appeared without supporting evidence. Students demonstrated bounded trust, valuing the readings while requesting instructor oversight. The study contributes design principles for responsible implementation. An emergent secondary finding is that the intervention also appeared to function as implicit training in critical engagement with probabilistic text: over the course of the semester, students learned to request checkable evidence and to recognize when claims exceeded available warrants.

CORRESPONDING AUTHOR:
Alexander M. Sidorkin
California State University Sacramento, United States
sidorkin@csus.edu

KEYWORDS:
Generative AI; dual tailoring; individualized instruction; higher education; instructional design; verification literacy; textbook alternatives; scaffolding; epistemic trust

TO CITE THIS ARTICLE:
Sidorkin, A. M. (2026). From One-Size Texts to Tailored Readings: Student Experiences with AI-Generated Course Materials. *Open Praxis*, *18*(3), pp. 506–522. DOI: https://doi.org/10.55982/openpraxis.18.3.1141

Open Praxis
DOI: 10.55982/openpraxis.18.3.1141

INTRODUCTION

Commercial textbooks remain a central infrastructure of higher education, yet they impose three structural constraints that are increasingly difficult to justify in professionally oriented graduate programs: cost, fixedness, and limited relevance. First, cost remains a persistent access barrier. Even as reported out-of-pocket spending has shifted downward due to rentals, subscriptions, and "day-one access" programs, students still face nontrivial course-material budgets, and many still forgo required materials when prices feel punitive (College Board, 2025; Florida Virtual Campus, 2022; National Association of College Stores, 2025). Large-scale survey evidence indicates that more than half of surveyed students reported not purchasing a required textbook due to cost, with downstream academic consequences that include taking fewer courses, not registering for a course, and earning poorer grades (Florida Virtual Campus, 2022). Second, conventional textbooks are fixed in level. They typically reflect a single "target reader," leaving instructors to compensate through supplementary notes, selective reading assignments, or remedial explanations. Third, conventional textbooks are fixed in relevance. They are designed for broad adoption, which is economically rational for publishers but pedagogically limiting for instructors trying to connect concepts to local contexts, student work settings, or the specific problem-framing that motivates adult learners.

The resulting instructional gap is well known: individualized readings have long been desirable, but operationally rare. The underlying reason is mundane and stubborn. Authoring time and version control scale poorly. An instructor can tailor explanations during discussion, but producing multiple written versions of weekly readings that vary by student interest and comprehension level is usually beyond what one person can sustainably maintain, even with teaching assistants. The empirical literature on individualized and differentiated instruction has repeatedly found that learning gains depend on interactions between learner characteristics and instructional conditions, but it also underscores how difficult it is to implement differentiation at scale with fidelity (Connor et al., 2016; Cronbach & Snow, 1977). Bloom's (1984) classic "2 sigma" argument is invoked here as a scalability framing, not as a claim about effect size: one-to-one tutoring can produce large gains, yet the delivery model is costly and hard to scale. This study does not test tutoring-level learning gains and does not include a comparison condition. In short, the idea of individualized instruction has been pedagogically compelling for decades, while the logistics of individualized written materials have remained prohibitive.

Generative AI changes the feasibility frontier. Large language models can draft readable instructional text rapidly, enabling instructors to produce course readings that can be revised, expanded, shortened, contextualized, and reframed with relatively low marginal time cost (Kasneci et al., 2023). This does not eliminate the old constraints, but it rearranges them. The primary bottleneck shifts from production capacity to quality assurance and instructional governance. Put plainly, generative systems can write fluent text that is inaccurate, biased, or overconfident, and they can fabricate citations or misattribute claims in ways that look academically plausible (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2023; Walters & Wilder, 2023). For this reason, this project treats AI-generated readings as a promising instructional instrument that requires explicit guardrails rather than blind optimism.

The intervention documented in this study was implemented in a graduate course in educational leadership (fictitious course name: LEAD 503). Instead of assigning a commercial textbook or a collection of pre-existing texts, the course used AI-generated weekly readings designed to meet two goals simultaneously. Goal one was coverage of core concepts expected in a survey course on educational leadership, including governance, decision-making, organizational communication, and HR. Goal two was dual tailoring (differentiation): adapting readings simultaneously to student interests (professional context, sector, examples) and to comprehension level (pacing, scaffolding, explanatory depth). This second dimension aligns with the broader individualization literature that treats readiness and prior knowledge as consequential moderators of instructional effectiveness (Connor et al., 2016; Cronbach & Snow, 1977).

The purpose of the study is to document the student-perceived value of this approach, not to claim universal effectiveness. The study uses a convergent mixed-methods design (Creswell,

2014) that combines survey data with systematic artifact analysis to identify visible markers of tailoring. The central claim is limited and practical: in at least one real graduate course, AI-generated weekly readings can function as an acceptable substitute for a commercial textbook when the instructor treats generation as draft production and uses explicit constraints to manage credibility risks (Kasneci et al., 2023; UNESCO, 2023).

Research questions:

- RQ1: How did graduate students perceive the usefulness, credibility, and usability of AI-generated weekly readings?
- RQ2: What forms of interest tailoring and comprehension-level tailoring are visible in the reading artifacts and in student commentary?
- RQ3: Under what conditions did students judge the AI-generated readings acceptable as a practical substitute for a commercial textbook in this course?

CONCEPTUAL FRAME AND RELATED LITERATURE

Individualized instruction and personalization have been advocated since at least the 1970s as ways to align teaching with learners' needs and interests. Recent scholarship distinguishes between **adaptive** systems, which automatically adjust content and pacing based on a student's performance, and **personalized** learning, which aims to craft experiences that are meaningful and connected to the learner's context (Hernández-Herrera, Ortiz-Bejar, & Ortiz-Bejar, 2026). Adaptive systems can modulate difficulty or recommend next steps, but they typically operate within one canonical body of material. Personalized learning aspires to go further: it positions students' backgrounds, goals, and curiosities as drivers of content selection. The shift from uniform materials to adaptive or personalized experiences has been enabled by advances in learning analytics and artificial intelligence, yet a persistent logistical barrier remains. Producing multiple versions of weekly readings, each pitched to different interests or levels, is more labor-intensive than adjusting problem sets or feedback; thus, readings in most courses remain one-size-fits-all.

Interest influences attention and persistence. Contextualizing problems in domains students care about triggers situational interest and increases perceived value (Priniski et al., 2018). Evidence from intelligent tutoring systems shows that personalizing content to students' out-of-school interests yields faster and more accurate solutions (Walkington & Bernacki, 2014), and context personalization increases situational interest especially for learners with low initial interest (Hogheim & Reber, 2015).

Tailoring to comprehension level involves adjusting readability, vocabulary load, scaffolding, and formative checks. Scaffolding refers to temporary support that reduces task complexity and fades as competence increases (Faber et al., 2024). Contingent scaffolding adapts support to learner needs; too much increases extraneous cognitive load, while too little fails to alleviate working memory demands (Faber et al., 2024). In text-based contexts, this means controlling vocabulary density, defining key terms, segmenting explanations, and inserting self-checking prompts.

Large language models offer a new tool for tailoring reading materials because they can generate fluent instructional text quickly. This capability shifts the bottleneck from production to quality assurance. Generative systems can create multiple versions of a reading with different examples or levels of detail, enabling dual tailoring along interest and comprehension dimensions. However, survey research shows that students are ambivalent about the outputs of generative AI. A recent study of engineering students in China found that more than half of respondents believed generative AI improved their learning efficiency, initiative, and creativity, but many expressed concerns about the accuracy and domain-specific reliability of AI-generated content (Fan, Deng, & Liu, 2025). The authors note that training data may include biases, inaccuracies, or outdated information, which can propagate into the generated text and mislead students. Although generative AI can support personalized learning and instant feedback, educators must teach students to verify information and to avoid overreliance on AI outputs (Fan et al., 2025). The broader implications of AI-generated content for open education have been explored in recent work on synthetic content and open educational resources, which

highlights both the potential for automatic content generation and the challenges of quality assurance and equitable access (Bozkurt, 2023). A collective manifesto on teaching in the age of generative AI further emphasizes that AI is not culturally or ideologically neutral and that its integration requires critical, evidence-based oversight (Bozkurt et al., 2024). Instructors who use AI to generate readings should combine retrieval from vetted sources, manual verification, and explicit prompts to the model (e.g., emphasizing citations and explanations) to manage these risks.

Epistemic trust is not guaranteed even in formal education. Corbitt et al. (2024) found that graduate students across disciplines detected misinformation and bias in course content and varied in how much they trusted instructor claims. Commercial textbooks carry their own risks of inaccuracy and single-perspective framing. AI-generated readings add risks of hallucination and weak citations but can be tuned to student interests and comprehension levels. When instructors curate, verify, and integrate class discussion around the material, the quality can be comparable to conventional textbooks for many purposes.

COURSE CONTEXT AND INTERVENTION DESIGN

COURSE AND PARTICIPANTS

This study was conducted in Fall 2025 in LEAD 503: Organizational Systems and Human Resources, a masters-level course in an educational leadership program. It is a fictitious identifier to reduce deductive disclosure while preserving the instructional design features relevant to the study. The course focused on examining, synthesizing, and applying human resources and organizational management practices necessary for leadership in educational settings. Learning outcomes emphasized applying organization theory to educational environments, cultivating safe and productive working and learning climates, conducting recruitment and selection processes, and developing skills for performance evaluation.

LEAD 503 was delivered in a hybrid format combining in-person and synchronous online sessions. The course integrated applied simulations, case-based discussion, and structured policy-analysis activities that provided repeated occasions for students to translate reading-based knowledge into decisions and written products.

The study's primary evidence about student perceptions comes from an end-of-course survey administered after course grades were submitted. For LEAD 503, 24 students completed the survey and consented to participation, including consent for analysis of AI-related assignments that documented interaction with the course assistant. The course enrollment may have been larger than the survey respondent count, but the analytic sample for this paper is limited to consenting survey respondents in LEAD 503.

CLASS COMPANION: PLATFORM AND CONFIGURATION

The Class Companion is a course-tutor bot built using OpenAI's Custom GPT feature within ChatGPT and configured to generate weekly readings, answer syllabus questions, test comprehension via scenarios, and provide assignment guidance while maintaining academic integrity. The system is explicitly framed as fallible ("powered by AI and can make mistakes") and is designed to remain course-focused and to defer to the instructor when uncertain. See Appendix 2 for bot behavior instructions.

The content-generation framework directs the bot to write extended readings (target 2,000 words) for working professionals in California higher education, emphasizing immediate applicability, explicit learning objectives, and structured elements such as reflection questions and practical exercises. The engagement style is interactive, with Socratic prompts and requests for examples from the student's work setting. The configuration specifies boundaries including redirection of off-topic requests and refusal to complete assignments directly. The full instruction set appears in Appendix 2.

ASSISTANT ROLE AND INTEGRATION WITH THE COURSE SCHEDULE

The intervention consisted of an AI-powered "Class Companion" embedded as a standing course resource. Students were instructed to use ChatGPT or an equivalent major AI platform,

and they were encouraged to set up access in a way that enabled consistent functionality across the term. The Class Companion was positioned as a built-in tutor and course guide. Its intended functions included: (a) generating weekly readings aligned to the syllabus sequence, (b) answering questions about course requirements and deadlines, (c) assessing student understanding through scenarios and questions, and (d) providing assignment support by suggesting approaches and strategies.

For the purposes of this paper, the focal design feature is the weekly reading generator. Weekly readings were not ancillary; they functioned as the core "text" for the course. Readings were aligned to the topical sequence: recruitment and selection; governance and decision-making structures; policy development and implementation; information systems and communication flows; employee relations; crisis management and organizational communication; stakeholder relations and community engagement; performance management and employee development; strategic communication and public relations; organizational change and systems improvement; and leadership and organizational culture in higher education. In class, students repeatedly used these readings as shared preparation for discussion and as a conceptual substrate for applied tasks. Early in the term, class time included explicit demonstrations of the reading workflow and the reading submission format, followed by small-group discussion activities structured around the AI-generated readings.

Readings were operationalized as "reading logs" submitted to the learning management system as PDFs. Students were instructed to request the week's topic, generate a reading, and ask at least three follow-up questions before submitting. This requirement treated reading as an interaction rather than a one-shot consumption activity. Weekly postings were assigned course points (3 points each, 33 points total) but were not formally graded for quality, functioning as a preparation mechanism rather than an evaluative essay.

DEFINITION AND OPERATIONALIZATION OF "DUAL TAILORING"

The course design emphasized a specific form of individualization that this study refers to as dual tailoring. The term is a descriptive design label, not a new theory. It names an instructional configuration that combines two well-established dimensions of individualization within a single AI-mediated workflow. Dual tailoring has two dimensions.

1. First, interest tailoring refers to adapting the reading to the learner's professional context, sector, and topical curiosities. In a graduate leadership course, "interest" is often expressed as a preference for a particular institutional setting (community college, CSU, UC, private nonprofit), a professional role (department chair, HR analyst, dean, student affairs administrator), or a problem type (hiring, conflict resolution, governance breakdowns, crisis response). Interest tailoring can be implemented through changes in examples, cases, and applied prompts without changing the conceptual backbone of the reading.
2. Second, comprehension-level tailoring refers to adjusting the accessibility of the reading to students' readiness and background knowledge. In practice, this can involve altering pacing, adding definitions and intermediate steps, reducing vocabulary density, providing additional scaffolding (advance organizers, summaries, check-for-understanding questions), or offering a deeper "second layer" when a student requests more nuance. Importantly, comprehension-level tailoring does not imply remediation. In graduate professional learning, it often means enabling students to calibrate depth and explanation to the cognitive demands of their work context and prior exposure to HR, policy, and organizational theory.

Both forms of tailoring were supported by the same mechanism: iterative prompting. Students could request a first-pass reading and then ask targeted follow-ups to reframe, deepen, simplify, localize, or apply. The instructor controlled key constraints through the reading prompt template and course norms, including expected length, required elements (theory plus applied examples plus reflection questions), and the expectation that students should engage in multiple follow-up queries rather than accept a single generated text.

INSTRUCTIONAL NORMS, VERIFICATION EXPECTATIONS, AND ACCOUNTABILITY

The course adopted an explicit AI use policy that normalized AI as part of the learning workflow while preserving student responsibility for quality, veracity, and originality. Verification norms were reinforced through specific assignments, including a "policy compliance bot" task requiring accuracy checks, and through the reading workflow itself, which expected students to request explicit topics and constraints, ask clarifying follow-ups, and treat readings as inputs to higher-control activities (case analysis, simulations, writing tasks) assessed by the instructor. These applied activities served as a quality-control layer where misconceptions surfaced through discussion, peer critique, and instructor feedback.

DATA SOURCES AND METHODS

DATA SOURCES

Two primary sources inform the analysis: (1) an end-of-course survey with closed-response items and open-ended responses, and (2) weekly reading logs submitted as PDFs. The survey supplies student-perceived value, acceptability, and risk perceptions. See the actual instrument in Appendix 1. The logs supply two distinct but linked artifact types produced in normal course activity: the AI-generated reading text and the student–AI interaction that generated, revised, and extended that text. The reading logs are 4,487 pages total, or about 939,000 words.

READING LOGS AS DUAL ARTIFACTS

Each weekly log is a combined record with two analyzable layers. The first layer is the generated "reading," a textbook-style narrative with headings, learning objectives, summary points, and applied exercises. The second layer is the interaction trace, typically alternating student prompts and Class Companion responses. This structure enables two kinds of claims: claims about properties of the generated instructional text itself (for example, sector anchoring, role framing, definitional density) and claims about adaptive behavior in response to student prompts (for example, elaboration after "explain more," re-scoping after "that is not my job," or requests for cases and sources).

A minority of PDFs contained generated text with little visible follow-up interaction; these segments are usable for artifact-level analyses but excluded from analyses requiring prompt-response adjacency.

The artifact corpus includes logs B through K. Log A is excluded because it used chat links rather than the standardized PDF format, and mixing capture formats would introduce avoidable measurement noise.

SURVEY SAMPLING AND MEASURES

Survey analysis is restricted to consenting respondents who completed the survey items tied to Class Companion use and perceived outcomes. Closed-response items are treated as descriptive measures of perceived utility and acceptability, rather than as latent constructs supporting strong inference. Open-ended responses are treated as qualitative evidence of perceived benefits, concerns, and self-reported norms for responsible use.

The survey was a purpose-built descriptive instrument designed for this specific course context, not a previously validated scale. Items were constructed to capture perceived utility, acceptability, and boundary awareness rather than to measure latent constructs. Internal consistency (Cronbach's alpha) is reported only as a diagnostic for aggregation where item blocks plausibly cohere. This approach is consistent with the study's exploratory, proof-of-principle scope.

QUANTITATIVE ANALYSIS

Quantitative reporting is descriptive. For closed-response items, analysis focuses on response distributions and summary statistics. Where blocks of items plausibly form a coherent index, internal consistency can be reported as a diagnostic (for example, Cronbach's alpha) to

justify aggregation for descriptive purposes. Exploratory contrasts by usage intensity are only appropriate if the survey includes an interpretable and non-sparse usage measure; when that condition is not met, usage-based subgroup claims are omitted.

QUALITATIVE ANALYSIS

Qualitative analysis uses a focused thematic approach guided by the paper's conceptual claims. Coding targets (a) perceived value and acceptability (clarity, preparedness, relevance, credibility), (b) risk awareness and skepticism, and (c) perceived boundaries such as reliance concerns or demands for instructor oversight. Code development follows an iterative process: an initial pass to generate candidate codes, consolidation into a compact codebook aligned with the Findings structure, and a confirmatory pass to ensure coverage and identify deviant cases. In a single-course pilot, transparency and traceability of excerpts matter more than elaborate coder infrastructure; agreement checks can be applied to a subset if multiple coders are used, but the primary emphasis is on auditability of claims through short, attributable excerpts.

ARTIFACT ANALYSIS

Artifact analysis addresses two adaptation dimensions that are structurally encouraged by the bot configuration: interest tailoring and comprehension-level tailoring. Interest tailoring is operationalized through indicators of sector anchoring, role framing, topic emphasis, and between-log differentiation. Comprehension-level tailoring is operationalized through triggered increases in scaffolding after student prompts that request explanation, definition, or elaboration. The detailed operational definitions and any sampling strategies used for tractability are reported in the corresponding Findings subsections to minimize redundancy. One element warrants advance notice: between-log differentiation is assessed in part through TF-IDF cosine similarity, a standard computational measure of vocabulary overlap between documents. Readers unfamiliar with this technique will find a full explanation alongside the results in Finding 2. Table 1 summarizes the operational definitions used for both dimensions (see Table 1).

Table 1 Operationalization of dual tailoring in artifact analysis.

DIMENSION	INDICATORS	DATA SOURCE	DETECTION METHOD
Interest tailoring	Sector anchoring (institution-type references); role framing (professional role labels); tailoring markers (second-person role cues, adaptation offers)	Reading logs B-K (1/5-page sample)	Dictionary-based automated counts, normalized per 10,000 words; TF-IDF cosine similarity across logs
Comprehension-level tailoring	Definitional scaffolding (explicit definitions, paraphrases); stepwise structures (numbered sequences); check-for-understanding cues	Three selected logs (B, C, I)	Manual extraction of comprehension-oriented prompts; scaffolding density markers per 1,000 words in triggered vs. baseline responses

INTEGRATION AND EXCERPT SELECTION

Integration uses joint displays and paired narrative evidence. Tables provide log-level indicators, while excerpts document mechanisms that counts alone cannot show. Excerpts are selected using maximum-variation logic, prioritizing short passages with unambiguous interpretive warrants.

ETHICS AND CONFIDENTIALITY

The study involves a small class setting and artifacts that can contain rich professional context. Deductive disclosure is therefore a primary risk. All excerpts are de-identified, unique workplace identifiers are removed or generalized, and log labels are used instead of student names. Where a passage is distinctive enough to plausibly identify a participant, the passage is omitted or paraphrased while retaining the analytic claim through alternative evidence. The study was reviewed and approved by the university's Institutional Review Board prior to data collection.

Data availability. The survey instrument is provided in Appendix 1. De-identified summary data from the survey are available from the corresponding author upon reasonable request. The full reading logs cannot be shared publicly because they contain professional context sufficient to enable deductive identification of participants in a small class setting.

FINDINGS

FINDING 1: STUDENT-PERCEIVED VALUE AND ACCEPTABILITY

Students generally treated the Class Companion as a practical support tool. Across five rated functions (4-point effectiveness scale), ratings clustered in the upper range (Table 2). The strongest perceived value was assignment support: 83% rated the tool "Effective" or "Very Effective" (M = 3.38, SD = 0.77). Providing information about course requirements was also rated highly (88% Effective/Very Effective). Generating quality reading materials was positive but more mixed (75% Effective/Very Effective), suggesting that students accepted the readings as "good enough" while noticing variability. The lowest ratings were for assessing knowledge and skills (71% Effective/Very Effective), the only function with any "Not effective" selections.

On broader acceptability items (4-point agreement scale), most respondents endorsed the overall learning value and future willingness to use a similar tool. Seventy-five percent agreed that they learned more than they would have in a comparable course without an AI companion (29% fully agree; M = 2.96, SD = 0.91). Seventy-one percent said they would take another course with an AI class companion or equivalent (29% fully agree; M = 2.92, SD = 0.93). The strongest endorsement was skill development: 88% agreed that their AI skills increased significantly, and two-thirds fully agreed (M = 3.50, SD = 0.83). These results suggest that acceptability was not only about convenience. Students experienced the intervention as an opportunity to build transferable competence in using AI systems for academic work, which likely contributed to their willingness to see the approach repeated.

Open-ended comments sharpen how students connected value to clarity, preparedness, relevance, and credibility. On clarity and preparedness, multiple respondents described the tool as a fast, responsive "second set of directions" that helped them interpret assignments and organize work. One student emphasized the value of targeted explanation, noting that it "helped explain assignments and textbook readings," and another framed it as a time-saver that reduced friction in week-to-week preparation. These remarks align with the stronger quantitative ratings for course requirements and assignment completion, and they reinforce an interpretation of the assistant as a just-in-time support layer rather than a substitute instructor.

Perceived relevance appeared in two forms. First, students valued the ability to tailor examples to their professional context and local setting. A respondent explicitly requested that the tool use "examples related to Sac State and the California education system," implying that relevance was not merely topical alignment with the syllabus but contextual specificity. Second, students appreciated the ability to adapt the reading experience to their needs, including summarizing, re-explaining, or extending content. At the same time, a few comments suggested diminishing returns when the tool expanded beyond what the student asked for, for example describing it as "wordy" and inclined to "keep adding information." That minor friction is consistent with the more mixed reading-material ratings: usefulness was common, but not uniform.

Perceived credibility was a conditional stance rather than a rejection. Students articulated that the tool was helpful but should not become an unquestioned authority. One wrote, "I do not think we should get used to using it for everything," while still calling it "a great tool for learning." Another wanted clearer sourcing, noting that citing authors directly would make it "more credible." These comments signal bounded trust: students accepted the tool for scaffolding and navigating tasks but remained alert to limits and wanted cues supporting verification.

Table 2 Student perceptions of Class Companion value and acceptability (n = 24).

Note. Percentages are based on n = 24 respondents. Means and SDs are computed on the 1–4 scales indicated in the panel headings.

PANEL A. EFFECTIVENESS RATINGS (1 = NOT EFFECTIVE, 4 = VERY EFFECTIVE)								
ITEM	n	MEAN	SD	NOT EFFECTIVE n (%)	SOMEWHAT EFF. n (%)	EFFECTIVE n (%)	VERY EFFECTIVE n (%)	EFF./VERY EFF. n (%)
Course requirements (information)	24	3.25	0.68	0 (0.0%)	3 (12.5%)	12 (50.0%)	9 (37.5%)	21 (87.5%)
Understanding course concepts	24	3.04	0.69	0 (0.0%)	5 (20.8%)	13 (54.2%)	6 (25.0%)	19 (79.2%)
Completing assignments	24	3.38	0.77	0 (0.0%)	4 (16.7%)	7 (29.2%)	13 (54.2%)	20 (83.3%)

(Contd.)

PANEL A. EFFECTIVENESS RATINGS (1 = NOT EFFECTIVE, 4 = VERY EFFECTIVE)								
ITEM	**n**	**MEAN**	**SD**	**NOT EFFECTIVE n (%)**	**SOMEWHAT EFF. n (%)**	**EFFECTIVE n (%)**	**VERY EFFECTIVE n (%)**	**EFF./VERY EFF. n (%)**
Generating quality reading materials	24	3.12	0.8	0 (0.0%)	6 (25.0%)	9 (37.5%)	9 (37.5%)	18 (75.0%)
Assessing knowledge and skills	24	2.88	0.8	1 (4.2%)	6 (25.0%)	12 (50.0%)	5 (20.8%)	17 (70.8%)
PANEL B. AGREEMENT ITEMS (1 = DISAGREE, 4 = FULLY AGREE)								
ITEM	**n**	**MEAN**	**SD**	**DISAGREE n (%)**	**SOMEWHAT DISAGR. n (%)**	**SOMEWHAT AGREE n (%)**	**FULLY AGREE n (%)**	**AGREE (SOMEWHAT+FULLY) n (%)**
Learned more than without AI companion	24	2.96	0.91	2 (8.3%)	4 (16.7%)	11 (45.8%)	7 (29.2%)	18 (75.0%)

FINDING 2. EVIDENCE OF INTEREST TAILORING

Interest tailoring was a recurring and observable feature of the AI-generated reading logs. It appeared both as overt customization moves (explicit prompts to adapt examples, cases, or exercises to the student's setting) and as embedded contextualization (institutional names, sector-specific constraints, and role-relevant tasks). Table 3 summarizes these indicators across reading logs B–K, using a systematic sampling and a set of operational markers that translate qualitative differences into comparable numeric signals.

Method and operationalization were designed to balance feasibility with defensible inference across a very large corpus. Rather than hand-coding the full set of logs, we used a systematic automated 1/5-page sample from each log, producing samples ranging from 13,728 to 22,034 words per log (Table 3). Within each sample we counted a dictionary-based set of "tailoring markers" and normalized them per 10,000 words. These markers included second-person role anchoring (for example, "in your role"), direct adaptation offers (for example, "Would you like me to tailor…"), and concrete situational prompts (for example, "describe your case"). We also counted explicit sector cues via institution-type references, including "community college," "CSU," and "Sacramento State," again normalized per 10,000 words (Table 3). Finally, to estimate how different the logs were from one another in overall language use, we computed TF-IDF vectors for each sample and calculated cosine similarity across logs, reporting each log's average similarity to the rest and its minimum and maximum similarity values (Table 3). Cosine similarity provides a compact estimate of textual proximity: higher values indicate that two samples rely on more similar vocabularies and phrasing patterns, while lower values indicate greater divergence in word choice and emphasis.

Three patterns in Table 3 support the claim that interest tailoring was not merely anecdotal, but structurally present and measurably variable. First, the overall rate of tailoring markers was substantial, with a mean of 52.24 markers per 10,000 words (SD = 13.73), and a wide observed range from 38.74 (Log K) to 74.29 (Log B) (Table 3). In plain terms, customization language was frequent, but not uniform. Some logs leaned heavily into prompts, personalization offers, and role anchoring, while others relied more on generic exposition.

Second, sector focus varied in ways consistent with the interest-tailoring hypothesis. Log B is the clearest community-college-centered case, with 33.43 community college references per 10,000 words (Table 3). In contrast, most other logs are framed as "CSU/regional public," showing higher Sacramento State reference rates (e.g., Log F at 9.53 per 10,000 words). The AI selected plausible organizational subunits and missions fitting each institutional context, not merely changing labels.

Third, role framing varied across logs. Some are anchored in HR and employee relations (Logs E and H), others in IT and information systems (Logs C and K), and several in communications

and public relations (Logs D, F, G, I, J) (Table 3). Logs often made role framing explicit through conversational repair: in Log F, the student rejects a compliance-focused action plan, triggering a reframing to a nonprofit student-success role.

Representative cases show these dimensions co-occurring. In crisis communication materials, the AI embeds sector variation by offering parallel examples differing by institution type, paired with role-relevant tasks. In stakeholder engagement materials, the AI switches to a UC system lens when prompted, expanding into institution-specific narratives driven by the student's selected case.

The similarity results help quantify "degree of differentiation" beyond marker counts. Average cosine similarity across logs ranges from 0.50 to 0.61 (Table 3). Minimum similarity values fall as low as 0.41 (Logs B and H), indicating that at least one pairing for those logs diverges notably in vocabulary and emphasis, while maximum similarities reach 0.78 (Logs F and I), indicating that some logs share substantial shared language patterns (Table 3). Interpreted conservatively, this pattern supports a hybrid structure: a common instructional backbone across logs, with meaningful contextual overlays that shift the center of gravity toward a given sector and role. The result resembles a template with adjustable dials. The template keeps the course-aligned scaffolds stable, while the dials tune examples, institutional references, and practical prompts toward the student's declared setting.

As a robustness check, we recomputed similarities under alternative preprocessing choices: with and without stopword removal, and using bigram rather than unigram TF-IDF vectors. The pattern of results remained stable across these alternatives. Average similarity values shifted by no more than 0.03, and the rank ordering of logs by average similarity did not change, supporting the conclusion that the differentiation pattern is not an artifact of a particular preprocessing decision.

Table 3 Indicators of interest tailoring across reading logs (B–K).

Note. Metrics computed from a systematic 1/5-page sample of each PDF (every 5th page). Similarity is TF-IDF cosine similarity across sampled text. Values range from 0 to 1; a value of 1.0 would indicate identical vocabulary distributions, while 0.0 would indicate no shared vocabulary. In this corpus, average similarities between 0.50 and 0.61 indicate that logs share a common instructional core but diverge substantially in sector-specific and role-specific language.

LOG	B	C	D	E	F	G	H	I	J	K
Pages in sample (1/5)	84	75	65	88	108	106	95	93	103	86
Words in sample	16153	16003	13728	17088	22034	21182	19937	18725	21090	18069
Tailoring markers per 10k words	74.29	66.86	54.63	71.4	45.38	39.66	44.14	47	40.3	38.74
Sacramento State refs per 10k	3.71	6.87	8.74	8.78	9.53	6.14	6.02	5.87	8.06	4.98
Community college refs per 10k	33.43	16.87	20.4	26.33	9.98	17.47	19.06	16.02	20.86	15.5
CSU refs per 10k	29.1	20.62	15.3	14.04	24.05	17.94	15.55	18.16	8.53	10.52
Avg cosine similarity to other logs	0.5	0.52	0.53	0.53	0.54	0.58	0.51	0.56	0.61	0.56
Min cosine similarity	0.41	0.47	0.43	0.46	0.44	0.5	0.41	0.44	0.56	0.48
Max cosine similarity	0.6	0.58	0.64	0.65	0.78	0.69	0.65	0.78	0.76	0.76

FINDING 3. EVIDENCE OF COMPREHENSION-LEVEL TAILORING

The reading logs contain repeated sequences where a student asks for clarification, expansion, or a definition, and the Class Companion responds by increasing scaffolding. In these episodes, scaffolding takes a recognizable form: explicit definitions, paraphrases ("in simple terms"), concrete examples, stepwise breakdowns, and occasional check-for-understanding prompts that ask the student to apply the idea to their local context.

Method

The corpus is large, so the analysis focused on three logs selected to span distinct topical and role contexts (B: governance and decision making; C: policy and change implementation with theory prompts; I: strategic communication and system versus campus framing). Each PDF includes embedded turn markers ("You said:" followed by "LEAD 503 Class Companion said:"), allowing extraction of student prompts and the immediate AI response. Prompts were flagged as comprehension-oriented when they explicitly requested explanation or conceptual clarification (for example, "what is," "explain," "elaborate," "in simple terms," "difference between"). For each extracted response, we computed a simple "scaffolding density" indicator

based on the rate of definitional and paraphrase phrases ("is defined as," "refers to," "in other words," "let us unpack," "in plain terms"), stepwise structures ("Step 1," ordinal sequences, numbered lists), and check-for-understanding cues ("quick check," "try this," "your turn," "before you move on," "what would you do"). Rates were normalized by word count to support comparisons across response lengths. See more in Appendix 3.

Quantitative signal of adjustment

Across the three logs, definitional scaffolding was substantially more frequent in responses that followed comprehension prompts than in baseline explanatory text in the same logs. Measured as definitional markers per 1,000 words in AI responses, comprehension-triggered responses showed 3.4x to 8.7x higher definitional density than non-triggered responses (Log B: 0.69 versus 0.17; Log I: 0.61 versus 0.07; Log C: 0.37 versus 0.11). Check-for-understanding cues were also more likely to appear after comprehension prompts in two of the three logs, although they were less frequent overall than stepwise formatting. These ratios matter because they support the central claim of comprehension-level tailoring without requiring judgments about correctness or instructional quality. When students signaled uncertainty or asked for conceptual expansion, the AI produced measurably more definitional and clarifying language.

Representative cases

One clear instance of comprehension-level tailoring appears in Log B when a student asks a definitional question that also implies a need for usable language in a hierarchical setting: "what is framing questions." The Companion responds by shifting from general exposition to a scaffolded mini-lesson. It first supplies an explicit definition under a heading ("What Are Framing Questions?"), then lists features of the construct in bullet form (what framing questions do in meetings, what priorities they surface, how they guide conversation), then provides multiple situational examples tied to a specific context (student housing decision making). Only after the definitional and example layer does the response move to procedural guidance, providing a "Sample Script" and a stepwise structure ("Step 1" and subsequent steps). The form of the response indicates adaptive depth. It does not merely restate the earlier material, but expands the conceptual layer and then builds a procedural layer on top of it, matching the prompt's implied need for both understanding and enactment.

A second case, in Log C, shows how a request to "elaborate more on the theories such as Kotter's model" triggers a dense, stepwise explanation that embeds definitions and concrete higher-education examples. The response begins by naming the theoretical tool and immediately reframes it as practical leverage for implementation. It then provides a numbered list of Kotter's eight steps, with each step coupled to a higher-education policy scenario (for example, accreditation-driven assessment policy revisions, governance coalition building across faculty senate and administration, communication routines through faculty meetings and staff orientations). The response structure performs comprehension support through sequencing. It breaks a complex model into a series of smaller units, labels each unit, and anchors each label in a plausible institutional example. This is a direct mechanism for adjusting to a learner who needs both conceptual clarity and a bridge to application.

Taken together, these cases show a consistent interactional pattern: student prompts that signal insufficient comprehension lead to responses that increase definitional density, strengthen sequencing, and occasionally introduce micro-diagnostics that solicit context. The "tailoring" is therefore not only topical or sector-based. It includes adjustment in explanation depth and scaffolding form in response to learner signals embedded in the dialogue.

FINDING 4: BOUNDARIES, FAILURE MODES, AND STUDENT VERIFICATION RESPONSES

Technical and content boundaries

Boundaries and failure modes were visible in the reading logs as a set of recurring reliability risks rather than a single dominant problem. The issues clustered around sourcing and citation practice, occasional over-specific institutional claims, uneven quality of reference signals, and modest but identifiable student skepticism about fit and accuracy.

Method

We treated each log (B–K) as an interaction transcript and ran two automated screens over the full corpus (4,487 pages): (a) page-level presence of citation signals (APA-style citations, URLs, DOIs, Wikipedia mentions), and (b) institution-specific assertive statements using "Sacramento State" or "Sac State" near obligation or factual verbs. Student skepticism was counted from turns explicitly challenging accuracy or role fit. Counts are conservative, depending on specific lexical markers and PDF text extraction.

Sourcing problems and missing traceability

A core boundary is that many readings present as authoritative while providing limited traceability. Across the 4,487 pages, only about 0.80% of pages contain APA-style in-text citations in the form "(Author, year)" or similar. URLs appear on about 0.53% of pages, and DOI strings are essentially absent. This means that even when the text uses academic tone, most claims remain difficult to audit from within the artifact itself. The practical implication is not that the readings are necessarily wrong, but that students and instructors have limited internal means to confirm them, especially when claims become institution-specific or time-sensitive.

Uneven citation quality, including mixed-source lists

Where citation signals do appear, they are not consistently curated. Two related patterns recur. First, some logs embed conventional scholarly anchors that look appropriate for a graduate course (for example, classic stakeholder or crisis communication citations such as "(Freeman, 1984)" and "(Coombs, 2007)"), which can increase perceived credibility even when a full reference list is not provided (Log G, p. 111; Log I, p. 42). Second, several logs include explicit Wikipedia references alongside peer-reviewed sources. In one instance, a "Key Sources" list ends with "Wikipedia: Change Management, Organizational Theory, Governance in Higher Education, Crisis Communication, Human Resource Management" (Log J, p. 19). Another log includes items formatted as "Governance in Higher Education – Wikipedia (2024)" and "Crisis Communication – Wikipedia (2024)" within a reference cluster (Log I, p. 171). Corpus-wide, Wikipedia is mentioned on about 1.65% of pages, with variation by log (roughly 0.62% to 3.73% of pages).

Wikipedia was intentionally included in the knowledge base for accessible background orientation, but when it appears alongside peer-reviewed citations without explicit differentiation, students must infer source quality from context rather than from labels.

Over-specific claims without supporting evidence

A second boundary is the tendency to produce highly specific institutional claims in a confident voice. About 1.03% of pages contain patterns pairing "Sacramento State" or "Sac State" with assertive policy verbs. Representative examples include claims about revised RTP policies and CSU Executive Orders that may be accurate, partially accurate, or inaccurate, but the artifact provides no verifiable citation. This is the most consequential failure mode in an instructional setting, because specificity creates a higher cost of being wrong than generic content.

Student skepticism and boundary-setting

Direct skepticism in the logs is sparse. Across 837 student turns in B through K, one turn is an explicit challenge to role fit and output accuracy: "this is inaccurate, that is not my job" (Log F). The low frequency should be interpreted cautiously: it may reflect generally acceptable outputs, or it may reflect students treating logs as completion artifacts rather than venues for adversarial checking.

Survey comments reinforce this pattern. Four of 24 respondents used language of dependence or reliance, and one explicitly named the need for "teacher oversight" to ensure accuracy. One described becoming "somewhat codependent on the AI for reassurance and structure." Taken together, the corpus suggests a bounded credibility profile: the Companion generates readings that feel course-aligned and context-aware, but its outputs often lack internal traceability, its citation practice is inconsistent, and it occasionally produces high-specificity institutional claims without supporting evidence.

STUDENT VERIFICATION BEHAVIORS

Student responses to these boundary conditions were visible in the interaction logs. Across 837 student turns extracted from logs B–K, 23 turns (about 2.7 percent) contained explicit risk-

aware moves using conservative markers (corrections of AI assumptions, requests for evidence or cases, statements of confusion functioning as comprehension checks). The distribution was uneven: two logs contained no marker-coded turns at all, while most logs clustered in a narrow band of roughly 2 to 4 percent. This pattern indicates that a critical stance toward AI output was present but not uniform, supporting the claim that such stance is learnable and improvable rather than an automatic byproduct of using AI-generated text.

The clearest evidence of verification is direct correction of AI assumptions. In Log B, the student writes, "that is inaccurate I work in student housing at UC Davis," triggering a reset and revised framing. In Log F, the same stance appears as role-bounding: "this is inaccurate, that is not my job." Both cases illustrate that the risk in personalization is not only factual error but category error: the assistant may generate plausible text fitting the wrong job function. The students interrupt that failure at the prompt level.

A second, distinct form of critical stance is the demand for concrete, externally anchored evidence. In Log I, the student asks, "can you please show me actual case studies?" The phrasing is instructive. The student does not merely request more examples; they request actual cases, which implicitly raises a standard of warrant and traceability beyond generic vignettes. In the broader context of AI-generated instructional text, this kind of request is one of the few observable behaviors that directly increases epistemic quality, because it pressures the system toward specificity that can be checked against institutional reports, peer-reviewed literature, or publicly available documentation.

A brief sequence from Log I illustrates the verification cycle concretely. The student prompts: "Can you please show me actual case studies?" The Companion responds with a narrative about a named university's crisis communication strategy, citing general circumstances but providing no document title, date, or retrievable source. The student follows up: "Can you give me the sources for these case studies?" The Companion then supplies a list mixing plausible-looking citations with at least one that could not be independently confirmed. This three-turn sequence captures the core dynamic: the student raises the evidentiary standard, the system attempts to comply, and the result is partially but not fully traceable, leaving the student in a position where further external verification is needed.

A third form is the explicit comprehension check. In Log E, the student writes, "I have no idea what you are talking about." Functionally, this is a safeguard against unearned assent: it forces re-explanation and stepwise progression, which is precisely the mechanism through which comprehension-level tailoring becomes observable (Faber et al., 2024).

Pedagogically, this suggests that "critical thinking toward probabilistic text" can be taught as a situated literacy rather than asserted as a general virtue. In this intervention, the teachable unit is the move, not the attitude: correcting a mismatch in sector or role; requesting a checkable case; asking for citations that can be traced; and stating confusion early enough to trigger re-scaffolding. The logs show that some students already performed these moves while others did not, which implies room for explicit instruction, modeling, and peer norming grounded in authentic excerpts from the course's own interaction archive.

DISCUSSION AND IMPLICATIONS

The findings support a limited but consequential claim: AI-generated weekly readings can function as a practical substitute for commercial textbooks in at least some graduate professional courses, provided the instructor treats generation as draft production, applies systematic review, and uses explicit verification norms. Students accepted the readings as useful preparation, recognized value in their adjustability, and developed transferable skills in working with probabilistic text. At the same time, the boundary conditions matter. The readings often lacked internal traceability, citation quality was inconsistent, and over-specific institutional claims occasionally appeared without supporting evidence. This pattern suggests that the relevant question is not whether to use AI-generated readings, but how to structure their use so that fluency does not become a substitute for warranted knowledge.

DUAL TAILORING AS A DISTINCTIVE INSTRUCTIONAL MECHANISM

The dual-tailoring framework addresses the longstanding gap between the pedagogical value of individualized instruction and the practical barriers to implementation at scale (Bloom, 1984; Cronbach & Snow, 1977). Prior personalization approaches typically adjusted along a single dimension, such as difficulty level in adaptive problem sets or topical examples in intelligent tutoring systems (Walkington & Bernacki, 2014). Fan et al. (2025) found that engineering students recognized efficiency gains from generative AI but expressed concerns about accuracy and domain-specific reliability, a pattern closely mirrored in this study's bounded trust findings. This intervention demonstrates simultaneous adjustment along both dimensions described earlier.

The evidence confirms that this tailoring was structural, not cosmetic. Readings shared a common instructional backbone but diverged meaningfully in vocabulary and emphasis, as confirmed by cosine similarity analysis. The result resembles a template with adjustable dials that tune examples and institutional references without abandoning course-aligned learning objectives. Scaffolding increased in response to student signals of confusion, delivered just-in-time rather than predetermined by the instructor.

What makes this mechanism distinctive is the pairing of generative capacity with iterative prompting. Students were required to ask follow-up questions, creating repeated opportunities to request reframing and test understanding. Reading became an interaction rather than a consumption activity, with implications for the visibility of comprehension gaps that would otherwise remain latent.

THE HIDDEN PEDAGOGY OF VERIFICATION

A finding that extends beyond the initial research questions is that the intervention functioned as an implicit training ground for critical engagement with probabilistic text. The corpus contains repeated instances of students correcting the assistant's assumptions, requesting concrete evidence, and refusing to proceed when explanations became unintelligible. These moves were not uniformly distributed across students, which creates instructional leverage. When some students enact verification behaviors in visible ways, instructors can make those moves explicit, name them as desirable practice, and build routines that transfer them to peers. This finding aligns with Corbitt et al.'s (2024) observation that graduate students vary in how much they trust even instructor-provided claims, and extends it by showing that AI-generated text can make verification a recurring, observable practice rather than an occasional disposition. The broader educational community has recognized this challenge: a recent collective manifesto on generative AI in education calls for critical engagement rather than passive acceptance of AI outputs (Bozkurt et al., 2024).

The primary risk of AI-generated text is not that it is always wrong but that fluency can mask uneven quality. The AI-generated reading model distributes verification responsibility: students must learn to treat generated text as provisional and to recognize when claims exceed the evidence provided. Survey data confirm that students adopted this stance. They valued the readings for clarity and preparedness but also expressed caution about dependence and requested instructor oversight. This is bounded trust, not blanket acceptance, and it is arguably a more productive epistemic posture than the unexamined authority that commercial textbooks often enjoy.

The intervention suggests replicable design principles. First, treat generation as draft production; the instructor must review outputs and correct weak citations. Second, maintain explicit norms requiring follow-up questions and verification. Third, use applied activities (case analysis, simulations) as the quality-control layer. Fourth, position the assistant as a fallible guide rather than an authoritative text. Fifth, build iterative refinement into the workflow so that outputs improve across weeks.

A broader hypothesis emerges from these findings. Commercial textbooks hide their limitations beneath consistent formatting and publisher prestige; AI-generated text makes quality variance visible because the source is explicitly probabilistic. If we take seriously the idea that students should learn to evaluate claims and seek evidence, then a reading format that requires

verification may prove more pedagogically productive than one that naturalizes acceptance through professional presentation. This remains a hypothesis, not a finding of this study.

LIMITATIONS AND BOUNDARIES OF THE CLAIMS

The study documents one particular way of introducing AI-generated readings in a single graduate course with a small sample and limited artifact capture. The findings cannot support claims about universal effectiveness or transferability to undergraduate courses, large lectures, or disciplines with different epistemological norms. Transferability to undergraduate or large-enrollment settings is untested and constitutes a priority for future research, given the differences in student autonomy, prior knowledge, and the instructor's capacity for individual oversight that characterize those contexts. The survey data are self-reported perceptions, which are consequential for acceptability but do not establish learning gains relative to a control condition. The study does not support causal claims about learning outcomes. The artifact analysis uses conservative marker-based counts and systematic sampling, which provide defensible lower-bound estimates but may undercount subtler forms of tailoring and critical engagement.

The intervention also relies on instructor time for prompt design, output review, and verification scaffolding. That time cost is lower than authoring multiple textbook-length documents from scratch, but it is not zero. Instructors without experience in prompt engineering or without institutional support for AI integration may find the approach difficult to implement with fidelity. The findings are therefore best interpreted as feasibility evidence rather than as an endorsement for wholesale replacement of textbooks without corresponding changes in instructional infrastructure and faculty development.

Finally, the course took place in Fall 2025, at a moment when generative AI capabilities and institutional norms around AI use were still evolving rapidly. What is feasible and acceptable today may shift as platforms change, as students develop more sophisticated strategies for using AI, and as institutions develop clearer policies about AI-generated content in academic work. The boundary conditions documented here should be treated as context-specific rather than universal.

IMPLICATIONS FOR FUTURE WORK

The next phase of this work addresses the most consequential limitation: the uneven quality and weak sourcing visible in the Fall 2025 logs. In the subsequent semester, the instructor used ChatGPT's Deep Research workflow to assemble a curated set of course-aligned reference materials with explicit sourcing, then incorporated those materials into the Class Companion's knowledge base. This redesign shifts the generation model from predominantly prompt-driven text creation toward retrieval-augmented generation anchored in vetted sources. The expectation is that this approach will reduce over-specific unsourced claims, improve citation quality, and increase traceability while preserving the dual-tailoring capacity that students valued.

Broader implications extend to AI-integrated curriculum design. If dual tailoring is viable in one graduate course, it may be viable in other professionally oriented programs where relevance to work context is a persistent challenge. Future work should examine scalability to larger courses, transferability across disciplines, and whether the bounded credibility stance students adopt toward AI-generated text influences their approach to other authoritative sources. The findings here do not prove that AI-generated readings are superior to commercial textbooks, but they provide enough evidence to take seriously the possibility that a flexible, context-aware reading format may be a better fit for how graduate professional learning actually works.

DATA ACCESSIBILITY STATEMENT

The survey instrument is provided in Appendix 1 of the manuscript. De-identified summary data from the survey are available from the corresponding author upon reasonable request. The full reading logs cannot be shared publicly because they contain professional context sufficient to enable deductive identification of participants in a small class setting.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Appendices.** Appendix 1 to 4. DOI: https://doi.org/10.55982/openpraxis.18.3.1141.s1

SUSTAINABLE DEVELOPMENT GOALS (SDGs)

This study is linked to the following SDG(s): Quality education (SDG 4).

ETHICS AND CONSENT

The study was reviewed and approved by the university's Institutional Review Board prior to data collection. Students completed an end-of-course survey and consented to participation, including consent for analysis of AI-related assignments and interaction logs. All excerpts were de-identified and generalized to prevent deductive disclosure.

COMPETING INTERESTS

The author has no competing interests to declare.

AUTHOR CONTRIBUTIONS (CRediT)

Alexander M. Sidorkin: Conceptualization, methodology, formal analysis, investigation, data curation, visualization, writing—original draft preparation, writing—review and editing. The author has read and agreed to the published version of the manuscript.

AUTHOR NOTES

This paper was reviewed, edited, and refined with the assistance of Anthropic's Claude (Claude Opus 4, as of March 2026), complementing the human editorial process. The human author critically assessed and validated the content to maintain academic rigor. The author also assessed and addressed potential biases inherent in the AI-generated content. The final version of the paper is the sole responsibility of the human author.

AUTHOR AFFILIATIONS

Alexander M. Sidorkin orcid.org/0000-0003-1083-8328
California State University Sacramento, United States

REFERENCES

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*(6), 4–16. https://doi.org/10.3102/0013189X013006004

Bozkurt, A. (2023). Generative AI, synthetic contents, open educational resources (OER), and open educational practices (OEP): A new front in the openness landscape. *Open Praxis*, *15*(3), 178–184. https://doi.org/10.55982/openpraxis.15.3.579

Bozkurt, A., Xiao, J., Farrow, R., Bai, J. Y. H., Nerantzi, C., Moore, S., Dron, J., Stracke, C. M., Singh, L., Crompton, H., Koutropoulos, A., Terentev, E., Pazurek, A., Nichols, M., Sidorkin, A. M., Costello, E., Watson, S., Mulligan, D., Honeychurch, S., … Asino, T. I. (2024). The manifesto for teaching and learning in a time of generative AI: A critical collective stance to better navigate the future. *Open Praxis*, *16*(4), 487–513. https://doi.org/10.55982/openpraxis.16.4.777

College Board. (2025). *Trends in college pricing and student aid 2025*. College Board.

Connor, C. M. D., Morrison, F. J., Fishman, B. J., Schatschneider, C., & **Underwood, P.** (2016). Individualizing student instruction in reading: Implications for policy and practice. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 54–61. https://doi.org/10.1177/2372732215624931

Corbitt, K., Hiltbrand, K., Coursen, M., Rodning, S., Smith, W. B., & **Mulvaney, D.** (2024). Credibility judgments in higher education: A mixed-methods approach to detecting misinformation from university instructors. *Education Sciences, 14*(8), 852. https://doi.org/10.3390/educsci14080852

Coombs, W. T. (2007). *Ongoing crisis communication: Planning, managing, and responding*. Sage.

Creswell, J. W. (2014). Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.). SAGE.

Cronbach, L. J., & **Snow, R. E.** (1977). Aptitudes and instructional methods: A handbook for research on interactions. Irvington.

Faber, T. J. E., Dankbaar, M. E. W., van den Broek, W. W., Bruinink, L. J., Hogeveen, M., & **van Merriënboer, J. J. G.** (2024). Effects of adaptive scaffolding on performance, cognitive load and engagement in game-based learning: A randomized controlled trial. *BMC Medical Education, 24*, 943. https://doi.org/10.1186/s12909-024-05698-3

Fan, L., Deng, K., & **Liu, F.** (2025). Educational impacts of generative artificial intelligence on learning and performance of engineering students in China. *Scientific Reports, 15*, 26521. https://doi.org/10.1038/s41598-025-06930-w

Florida Virtual Campus. (2022). *2022 student textbook and instructional materials survey: Results and findings* (Office of Distance Learning & Student Services). Tallahassee, FL.

Freeman, R. B. (1984). Longitudinal analyses of the effects of trade unions. *Journal of labor Economics*, *2*(1), 1–26. https://doi.org/10.1086/298021

Hernández-Herrera, J. R., Ortiz-Bejar, J., & **Ortiz-Bejar, J.** (2026). Adaptive and personalized learning in higher education: An artificial intelligence-based approach. *Education Sciences, 16*(1), 109. https://doi.org/10.3390/educsci16010109

Hogheim, S., & **Reber, R.** (2015). Supporting interest of middle school students in mathematics through context personalization and example choice. *Contemporary Educational Psychology, 42*, 17–25. https://doi.org/10.1016/j.cedpsych.2015.03.006

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

National Association of College Stores. (2025, August 20). *NACS Student Watch Report: Course materials spending stable, high satisfaction with access programs.*

Priniski, S. J., Hecht, C. A., & **Harackiewicz, J. M.** (2018). Making learning personally meaningful: A new framework for relevance research. *Journal of Experimental Education, 86*(1), 11–29. https://doi.org/10.1080/00220973.2017.1380589

UNESCO. (2023). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization.

Walkington, C., & **Bernacki, M. L.** (2014). *Motivating students by "personalizing" learning around individual interests: A consideration of theory, design, and implementation issues*. https://doi.org/10.1108/S0749-742320140000018004

Walters, W. H., & **Wilder, E. I.** (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports, 13*, 14016. https://doi.org/10.1038/s41598-023-41032-5

TO CITE THIS ARTICLE:
Sidorkin, A. M. (2026). From One-Size Texts to Tailored Readings: Student Experiences with AI-Generated Course Materials. *Open Praxis*, *18*(3), pp. 506–522. DOI: https://doi.org/10.55982/openpraxis.18.3.1141

Submitted: 25 January 2026
Accepted: 10 April 2026
Published: 04 August 2026

Open Praxis is a peer-reviewed open access journal published by International Council for Open and Distance Education.