Computers and Education Open 10 (2026) 100350

Contents lists available at ScienceDirect

Computers and Education Open

journal homepage: www.sciencedirect.com/journal/computers-and-education-open

The collision of pedagogy and algorithm: Emergence of a new gender gap?

Drake Mullens [*,a], Stella Shen [b]

[a] Tarleton State University, United States
[b] University of Texas at Tyler, United States

ARTICLE INFO

Keywords:
Human-AI collaboration
Higher education pedagogy
Gender equity
AI literacy
Career mobility
Skill displacement

ABSTRACT

Despite near-universal AI adoption among university students, no established framework connects student AI interaction patterns to the career outcomes they may predict, leaving institutions without the evidence to guide integration or assess the cost of inaction. This study establishes a preliminary empirical bridge between observed student engagement and occupational outcomes through a novel multi-rater protocol combining human expertise with architecturally diverse large language models. We conduct a macro-structural analysis of discipline-level patterns across the National Center for Education Statistics field of study taxonomy by mapping aggregated data from 574,740 student interactions onto a validated career-competency framework. The findings reveal a structural alignment between pedagogical practice and economic stratification. Disciplines with the highest female representation occupy the lowest range of AI Interaction Scores. These same disciplines already face substantial earnings disadvantages, with a $52,090 gap separating the highest and lowest-earning fields. AI Interaction Scores correlate negatively with female representation ($\rho = -0.68$) and positively with median earnings ($\rho = 0.72$), suggesting unguided integration may entrench rather than ameliorate existing inequities.

1. Introduction

The integration of artificial intelligence into students' academic lives is no longer an emerging trend but a global norm. Recent large-scale surveys demonstrate near-universal adoption: 86 % of students globally now use AI in their studies [1], with another survey of over 11,000 undergraduates finding a similar rate of 80 % [2]. Moreover, the pace of this integration is accelerating rapidly, with student AI use surging from 66 % to 92 % in a single year [3]. This rapid assimilation establishes the urgent context for this study; yet, no pedagogical consensus exists on how integration should proceed.

The field lacks shared frameworks for what AI systems can reliably do, how students should engage them for learning versus efficiency, and which interaction patterns translate to workplace competency [4]. The invocation of 'AI literacy' as a curricular solution exemplifies the depth of this void. Literacy denotes foundational conceptual understanding, not the sophisticated human-AI collaboration increasingly demanded by labor markets [5]. Meanwhile, the disconnect between technical proficiency and effective workplace application persists, as many universities have yet to integrate AI comprehensively into their curricula [6,7]. This conceptual vacuum may disadvantage students' career prospects.

The integration of artificial intelligence into global economic systems represents not an incremental technological shift but a fundamental restructuring of work, skill valuation, and career trajectories [8–10]. As AI systems transition from specialized research tools to ubiquitous academic companions [11], educational institutions bear a dual obligation: cultivating disciplinary knowledge and fostering the human-AI competencies now essential for workforce participation [12,13]. Yet, the pedagogical paradigms and institutional policies governing student AI adoption lack empirical or theoretical linkage to the interaction patterns associated with upward career mobility [14,15]. This disconnect is not a benign oversight; it risks mis-calibrating educational investments toward AI literacy while neglecting the competency development that labor markets reward. This study addresses this lacuna by systematically mapping observed student AI interactions onto empirically validated career competency frameworks, revealing both promising opportunities and persistent inequities that suggest pedagogical intervention is needed.

Notably, this disconnect may compound existing inequality. Academic disciplines are stratified by gender (NCES, 2023a), and critical algorithmic scholarship has established that technology adoption without intentional equity-focused design perpetuates structural disparities [16,17]. If prevailing AI usage patterns differ across disciplines and align with the career-advancing versus automation-susceptible

* Corresponding author.
E-mail address: mullens@tarleton.edu (D. Mullens).

https://doi.org/10.1016/j.caeo.2026.100350
Received 18 June 2025; Received in revised form 3 February 2026; Accepted 21 March 2026
Available online 22 March 2026

distinction identified in occupational research, unguided integration risks encoding disadvantage not in the algorithms but in the interaction patterns that disciplines pedagogically promote or tacitly permit. Compounding documented technological biases and existing gendered economic stratification, we submit an additional vector of inequity rooted in pedagogical practice itself.

Beyond long-term career trajectories, the immediate learning environment itself misses significant opportunities when AI pedagogical gaps persist. Without targeted instructional frameworks guiding how AI tools are employed, students risk defaulting to superficial interactions characterized by cognitive offloading by soliciting direct answers or completed outputs with minimal personal engagement or critical reflection [18,19]. This interaction dynamic weakens foundational cognitive skills such as problem-solving, analytical reasoning, and information synthesis, undermining precisely the competencies higher education strives to cultivate [20,21]. Furthermore, disciplines lacking structured AI integration inadvertently foster passive learning environments rather than active, cognitively engaged classrooms. In such contexts, students lose opportunities to develop nuanced human-AI collaborative skills. These are skills essential not only for future professional success but also for immediate academic growth and intellectual maturity [22]. Addressing this pedagogical gap thus serves dual purposes: it safeguards immediate educational quality and lays the necessary groundwork for long-term career readiness.

To confront the aforementioned challenges, we draw on sociocultural learning theory's Zone of Proximal Development and critical perspectives on AI's societal impact, employing the AI-Accentuated Career Transitions (2ACT) framework [14]. Our study aims to:

1. Systematically map student AI interaction patterns onto the 2ACT framework to translate observed student behaviors into career-relevant AI interaction categories.
2. Quantify the overall prevalence of career-relevant (augmentation-focused vs. automation-focused) 2ACT patterns within observed student AI usage.
3. Examine associations between discipline-level AI Interaction Scores, median graduate earnings, and gender composition across the federal NCES field-of-study taxonomy.
4. Delineate specific, empirically informed, and theoretically grounded pedagogical recommendations designed to cultivate equitable, career-advancing human-AI collaborative skills.

This investigation contributes a macro-structural analysis of the complete NCES field-of-study taxonomy (574,740 conversations across 9 disciplines), identifying structural patterns that individual-level studies cannot detect to generate testable hypotheses for targeted primary research. The findings reveal statistically significant associations between AI interaction patterns, median graduate earnings, and gender composition, while the resultant pedagogical recommendations advocate targeted, equity-focused strategies amenable to intervention research.

The balance of work unfolds as follows: Section 2 reviews the relevant literature on AI's impact on occupational transformation, student AI engagement, and emerging equity disparities. Section 3 details the mixed-methods approach, including the conceptual mapping protocol and the derivation of the AI Interaction Score. Section 4 presents the findings from the conceptual mapping, the calculation of the overall 2ACT pattern prevalence, and the analyses. Section 5 discusses these findings, interpreting their alignment (and misalignment) with career-advancing competencies, critically examining the gendered implications, and connecting them to broader debates on skill development versus cognitive offloading. The articulation of detailed pedagogical recommendations is presented in Section 6. Finally, Section 7 outlines the study's limitations and proposes directions for future research, and Section 8 concludes with the overarching implications of this work for higher education in the AI era.

2. Literature review

The confluence of artificial intelligence and labor market dynamics has fundamentally altered the landscape of work and the competencies required for career advancement. To establish the need for evidence-based pedagogical frameworks that bridge current student AI usage with career-advancing competencies, this review synthesizes research across three domains: AI's transformation of work, emerging student AI engagement patterns, and the equity implications of this technological shift.

2.1. AI and the future of work: from automation to augmentation

The discourse on AI's labor market impact has matured from initial narratives of widespread job displacement [23] to more nuanced understandings of task transformation and human-AI complementarity [8,24]. While early research focused on technology's role in automating routine tasks and complementing non-routine cognitive work [24,25], recent scholarship reveals how AI specifically reshapes occupational tasks, creating new demands for human-AI collaboration [26,27].

This evolution reflects the changing nature of skill-biased technological change [28], where AI differentially impacts roles and potentially amplifies inequalities if access to requisite skill development remains uneven [29]. To effectively navigate this transformation, we must first understand which human-AI interaction patterns foster career advancement versus those that lead to skill displacement. While early scholarship foregrounded AI’s economic and occupational implications, recent pedagogical research emphasizes that successful integration of AI must prioritize immediate learning processes. This integration involves fostering students’ critical interactions with AI, thereby deepening active learning and promoting meaningful cognitive engagement within the classroom itself [30].

The 2ACT framework [14] empirically grounds this distinction. Developed through multivariate modeling of 545 occupations, 2ACT identifies two categories of human-AI usage patterns with starkly different career implications. Automation-focused patterns: Directive AI (where AI performs tasks with minimal human involvement) and Feedback Loop AI (where AI receives environmental feedback to complete tasks) both correlate strongly with lower occupational positioning. Conversely, augmentation-focused patterns: Validation AI (verifying human outputs), Task Iteration AI (collaborative refinement), and Learning AI (knowledge provision) are associated with higher job zones, particularly when combined with complementary human knowledge, skills, and abilities. The framework's emphasis on "Thinking Fraction" underscores that mere AI adoption is insufficient. How humans engage with AI determines whether technology becomes a ladder for advancement or a ceiling on mobility. This distinction becomes stark when examining how students currently interact with AI tools.

2.2. Student AI engagement patterns: promise and peril

Large-scale analyses of student AI interactions reveal both encouraging possibilities and career-blunting tendencies. Handa et al.'s [11] analysis of over half a million student conversations with Claude.ai provides unprecedented insight into real-world usage patterns. Their identification of four primary interaction modes: Direct Problem Solving (DPS), Direct Output Creation (DOC), Collaborative Problem Solving (CPS), and Collaborative Output Creation (COC), occurring with roughly equal frequency (23–29 %) suggests diverse engagement strategies.

The prevailing student AI engagement patterns identified by Handa et al. [11] carry significant implications for classroom learning environments. Expedient modes such as DPS and DOC, while efficient, reinforce shallow cognitive habits and hinder deeper disciplinary learning. Conversely, collaborative modes (CPS, COC) enrich classroom dialogue, foster metacognitive reflection, and elevate immediate learning outcomes through iterative refinement and collaborative

critique [19].

However, disciplinary variations in adoption and interaction styles raise important questions. STEM fields like Computer Science show early adoption and overrepresentation in usage, while Business and Humanities show lower adoption relative to enrollment [11]. Strikingly, nearly half (47 %) of student-AI conversations involve "Direct" interactions seeking answers or content with minimal cognitive engagement, thereby potentially fostering cognitive offloading rather than skill development.

From a sociocultural learning perspective [31,32], AI tools function as powerful mediational artifacts within students' learning ecologies. Effective learning depends not merely on tool availability but on the social context, task nature, pedagogical guidance, and student agency in interaction. Current patterns suggest that without robust scaffolding, students default to expedient modes (DPS, DOC) over deeper collaborative engagement (CPS, COC) that develops higher-order thinking. This gap widens in contexts where AI use is prohibited outright, driving usage underground and eliminating opportunities for guided skill development [33,34]. Without robust institutional scaffolding, students default to AI as a cognitive substitute rather than a cognitive partner [18], producing measurable consequences.

The risks are not theoretical speculation. Gillespie et al. [35] report that 77 % of students feel unable to complete coursework without AI assistance, while 81 % admit to reduced effort knowing AI is available. Over 75 % report relying on AI to perform tasks rather than learning to do them independently, and 66 % accept AI output without evaluating accuracy.

2.3. The equity dimension: gender, AI, and education

The integration of AI into education intersects with persistent gender disparities in academic pathways and career outcomes, yielding new forms of digital inequality ([36]; NCES, 2023a). Critical AI scholars and feminist technologists warn that without intentional equity-focused implementation, AI risks amplifying existing disparities [37,16,17].

Three factors converge to potentially amplify gender disparities. First, gender segregation in academic majors produces differential exposure to AI tools and pedagogical approaches. Fields with high AI adoption, like Computer Science and Engineering, traditionally have lower female representation (NCES, 2023a), while female-dominated fields (Education, Health Professions, Humanities) show different adoption patterns and interaction modes [11]. Female-dominated fields may lack structured access to meaningful, interactive, and reflective AI experiences, thereby limiting students' immediate opportunities and entrenching disparities in cognitive skill development [38].

Second, emerging evidence suggests institutional policy itself is gendered. Freeman [3] reports that male students are 45 % more likely to receive institutional encouragement to use AI, while female students are 38 % more likely to face prohibitions. This policy asymmetry creates what Freeman terms a widening "digital divide," where AI is simultaneously framed as academic misconduct in some contexts and a necessary career skill in others. Students navigating stricter regulatory environments may default to covert, transactional AI use rather than the overt skill-building interactions that develop career-relevant competencies.

Third, if career-advancing AI interaction patterns are more prevalent or valued in male-dominated fields while less advantageous patterns become normalized in female-dominated disciplines, AI integration could inadvertently fortify gendered economic disparities [39]. This structural dynamic transcends individual capability, highlighting systemic factors in how AI pedagogy is implemented across disciplines. Although not the focus of this study, AI tools themselves may embody biases that differentially impact students [40], compounding the challenge.

Despite these inequities, AI need not be disadvantageous. Targeted implementations have achieved equity gains in specific instructional contexts [41]. Optimism is warranted, though the efficacy of competency frameworks remains empirically untested.

2.4. The research gap: bridging student AI use with equitable career competencies

Despite parallel advancements in understanding AI's workforce impact [14], student AI engagement [11], and AI equity considerations [17], the pedagogical foundation remains underdeveloped. Gašević et al. [18] observe that design models for human-AI interaction in education have not yet been developed, leaving institutions without frameworks to guide the coordination between human cognition and AI capability. The field lacks both descriptive knowledge (how students actually interact with AI) and prescriptive frameworks (how they should). Handa et al. [11] provide a rare empirical window into the former; no systematic effort has connected these patterns to career-relevant outcomes that might inform the latter. This gap is obscured by a preoccupation with AI literacy as an educational objective. The distinction between competency and literacy is not mere pedantry: literacy focuses on knowledge and understanding, while competency involves applying that knowledge effectively and beneficially [5].

This disconnect is particularly pronounced at the disciplinary level, where gender demographics and pedagogical norms vary substantially. While the 2ACT framework delineates effective human-AI collaboration strategies for occupational mobility, and Handa et al.'s [11] work illuminates current student practices, these streams remain unconnected. Consequently, educational guidance remains uninformed. Without a theoretical grounding that links academic AI interaction to professional AI competency, targeted interventions cannot be designed. The absence of such interventions perpetuates a usage divide [3], potentially more consequential than digital divides of the past. AI pedagogy that fosters AI competencies is necessary to ensure that those without external access or privilege can also develop the strategic skills necessary for the AI economy.

3. Methodology

This study employed a phased, mixed-methods research design [42] to bridge empirically observed student AI interaction patterns with career-relevant AI competencies and examine equity implications. The design integrated qualitative conceptual framework mapping, systematic protocol validation, quantitative score derivation, and analysis of secondary data. The research pursued three aims: (1) address a theoretical gap linking educational AI usage to disciplinary labor market outcomes, (2) develop and validate a reusable protocol for LLM-assisted qualitative coding, and (3) derive discipline-specific AI Interaction Scores to examine associations with external indicators.

The study proceeded in three phases. Phase 1 established the conceptual mapping protocol through two stages: Phase 1a developed a baseline multi-rater protocol wherein four independent raters (one human expert, three large language models) translated student AI interaction patterns into career-relevant categories; Phase 1b validated this protocol through systematic variation of model architecture, generation, inference mode, and temperature, establishing boundary conditions for reliable LLM-assisted, producing evidence towards reusable methodological frameworks. Phase 2 applied the validated consensus mapping to derive discipline-specific AI Interaction Scores. Phase 3 examined associations between derived scores and external indicators of labor market outcomes and gender composition. Fig. 1 provides a visual overview of this design.

A note on analytic scope: the nine academic disciplines examined are the aggregated field of study categories used in NCES Digest of Education Statistics reporting (Table 318.20), excluding the residual ‘Other fields' category and account for 95.4 % of the Handa et al. [11] dataset. These disciplines represent the analytic population rather than a sample.

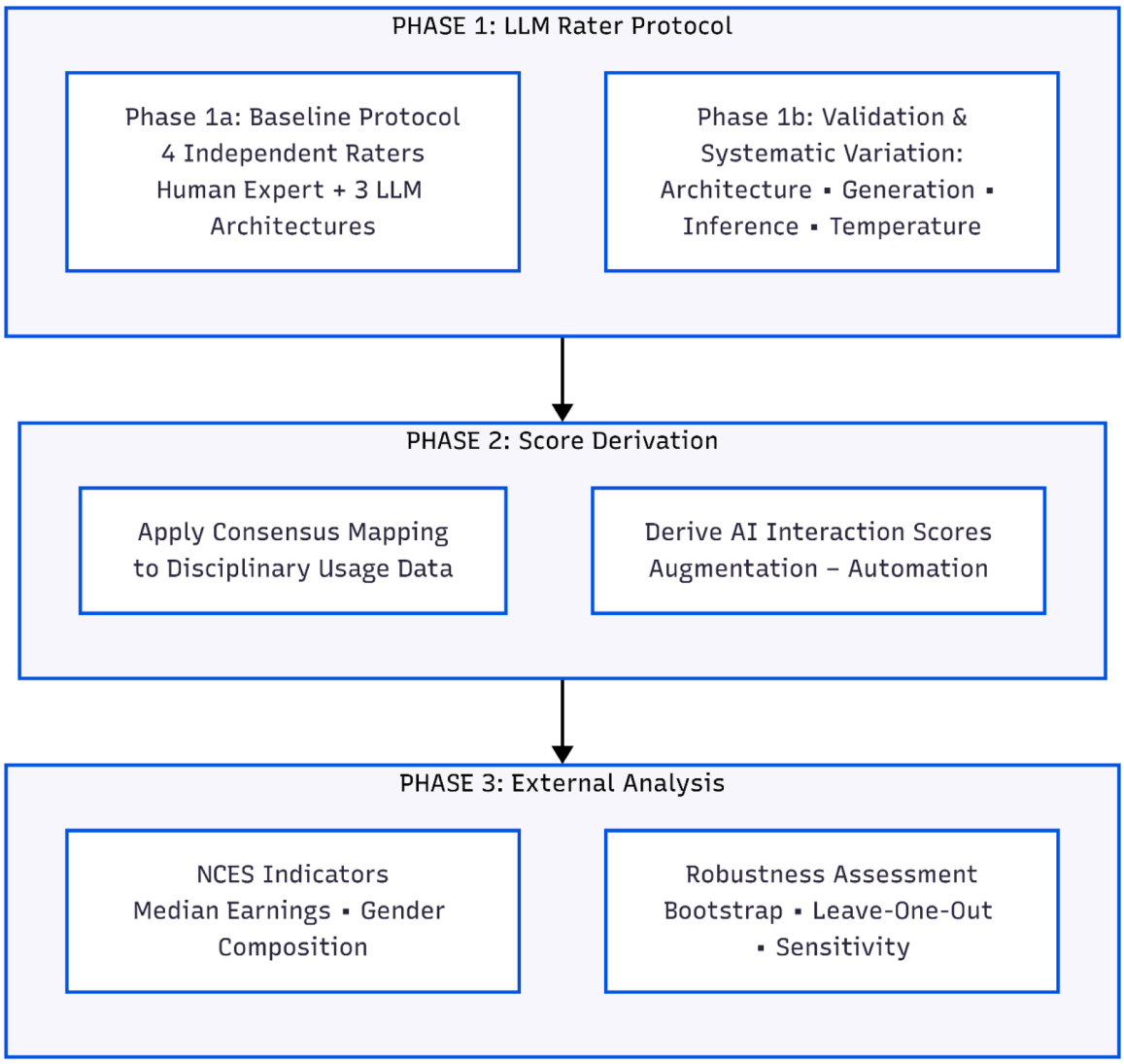

Fig. 1. *Methodological Flow for Mapping Student AI Usage to Career-Relevant Patterns*
Note. Phase 1 developed and validated a multi-rater protocol combining human expertise with architecturally diverse LLMs. Phase 2 applied the validated consensus mapping to Handa et al. [11] disciplinary usage data to derive AI Interaction Scores. Phase 3 examined associations between derived scores and NCES indicators of labor market outcomes and gender composition.

Consequently, statistical analyses describe relationships within this complete taxonomy rather than generalization from a sample to a population.

3.1. Phase 1: conceptual mapping and protocol development

This phase established the conceptual and empirical groundwork for the study, developed a multi-rater protocol for mapping student AI interaction patterns to career-relevant categories, and validated that protocol through systematic variation across model architectures, generations, and inference conditions.

3.1.1. Foundational frameworks

Student AI Interaction Patterns. Handa et al. [11] analyzed 574,740 anonymized student conversations with the Claude.ai model, identifying four primary interaction patterns: Direct Problem Solving (DPS), wherein students seek immediate answers or solutions; Direct Output Creation (DOC), wherein students request AI-generated content for direct use; Collaborative Problem Solving (CPS), wherein students engage AI as a thinking partner through iterative dialogue; and Collaborative Output Creation (COC), wherein students co-construct outputs through sustained human-AI collaboration. The authors reported the distribution of these patterns across nine academic disciplines, providing the empirical foundation for discipline-level analysis.

Career-Relevant AI Interaction Patterns. The AI-Accentuated Career Transitions (2ACT) Framework [14], developed through multivariate modeling of 545 occupations (O*NET Version 26.0), identifies five human-AI interaction patterns at the occupational level. Two patterns, Directive AI and Feedback Loop AI, are empirically associated with automation-focused work and lower occupational job zone placement. Three patterns, Validation AI, Task Iteration AI, and Learning AI, are associated with augmentation-focused work and higher job zone placement.

Structural Homology. Although these frameworks describe human-AI interaction in different domains and employ different disciplinary vocabularies, they share a fundamental structural correspondence. Handa et al.'s distinction between "direct" patterns (DPS, DOC) and "collaborative" patterns (CPS, COC) parallels the 2ACT framework's distinction between automation-focused patterns and augmentation-focused patterns. The underlying dimension is consistent: the degree of sustained human cognitive engagement throughout the task. Direct and automation-focused patterns involve delegation with minimal

iteration; the student or worker submits a task and receives a finished output. Collaborative and augmentation-focused patterns involve sustained human-AI exchange; the student or worker engages AI as a partner in iterative refinement, verification, or learning.

This structural homology provides the conceptual foundation for cross-domain mapping. By translating educational AI usage into the occupational vocabulary of the 2ACT framework, we can assess whether disciplinary interaction patterns align with career-advancing or automation-susceptible modes of engagement. The claim is not that educational and occupational contexts are identical, but that the structure of human-AI interaction exhibits sufficient cross-domain regularity that engagement patterns in learning contexts may prefigure engagement patterns in professional contexts.

Nevertheless, this mapping is inherently interpretive: no direct empirical link exists between how a student uses AI academically and how that usage corresponds to occupational categories. We therefore operationalized the mapping as a structured rating task amenable to multi-rater reliability assessment.

3.1.2. Phase 1a: baseline conceptual mapping protocol

The baseline protocol established the consensus mapping used throughout this study. The design prioritized accessibility and replicability, employing provider web interfaces under default settings to emulate a standard researcher workflow that presumes no programming expertise.

- **Rater Composition.** Four independent raters completed the mapping task: the primary researcher and three large language models representing distinct architectural lineages (OpenAI ChatGPT-o3 Pro, Anthropic Claude Sonnet 3.7 Extended Thinking, and Google Gemini 2.5 Pro Preview 05–06). These models were selected as state-of-the-art frontier models at the time of initial data collection. The use of three architecturally independent models from competing organizations guarded against the possibility that any single training corpus, reinforcement learning regime, or organizational alignment philosophy would dominate the consensus.
- **Temporal Independence.** The primary researcher completed all ratings prior to LLM engagement, ensuring that human judgment was not influenced by model outputs.
- **Task Definition.** Each rater independently assigned percentage allocations from each of the four student patterns to the five 2ACT patterns, with allocations for each student pattern constrained to sum to 100 %. This compositional structure required raters to make explicit trade-offs across categories rather than rating each independently.
- **Materials.** Raters received identical, verbatim textual descriptions of the four student patterns [11] and the five 2ACT patterns [14]. No additional context, examples, or anchoring information was provided, ensuring that ratings reflected conceptual alignment based solely on framework definitions. The complete prompt is reproduced in Appendix A.
- **Inter-Rater Reliability.** Intraclass Correlation Coefficients were computed using ICC(3,k), a two-way mixed-effects model assessing absolute agreement across average measures [43,44]. The analysis yielded: DPS, ICC = 0.974; DOC, ICC = 0.984; CPS, ICC = 0.962; COC, ICC = 0.949 (all $p < .0001$). All coefficients exceed the 0.90 threshold for excellent reliability [45].
- **Consensus Mapping.** Final allocations were derived by calculating the unweighted arithmetic mean across all four raters for each student-pattern-to-2ACT-pattern link, treating each rater as an exchangeable contributor to pooled judgment. Table 1 presents the resulting consensus mapping.

Table 1
Final consensus mapping of student AI interaction patterns to 2ACT career-relevant patterns.

Student AI Pattern	2ACT Pattern	Consensus Allocation (%)
Direct Problem Solving (DPS)	Directive AI	46.3
	Feedback Loop AI	2.5
	Validation AI	2.5
	Task Iteration AI	0
	Learning AI	48.8
	Total	100
Direct Output Creation (DOC)	Directive AI	76.3
	Feedback Loop AI	6.3
	Validation AI	3.8
	Task Iteration AI	3.8
	Learning AI	10
	Total	100
Collab Problem Solving (CPS)	Directive AI	1.3
	Feedback Loop AI	6.3
	Validation AI	8.8
	Task Iteration AI	41.3
	Learning AI	42.5
	Total	100
Collab Output Creation (COC)	Directive AI	1.3
	Feedback Loop AI	5
	Validation AI	18.8
	Task Iteration AI	62.5
	Learning AI	12.5
	Total	100

Note. The consensus mapping reveals an interpretable structure. Direct Output Creation loads heavily onto Directive AI (76.3 %), consistent with its characterization as delegation with minimal iterative engagement. Collaborative Problem Solving is distributed primarily across Task Iteration AI (41.3 %) and Learning AI (42.5 %), reflecting the dialogic, skill-building nature of this mode. Direct Problem Solving exhibits a bimodal allocation between Directive AI (46.3 %) and Learning AI (48.8 %), capturing its dual potential for either shortcut-seeking or genuine inquiry. Collaborative Output Creation loads predominantly onto Task Iteration AI (62.5 %), consistent with its emphasis on sustained co-construction. Percentages may not total 100.0 % due to rounding.

3.1.3. Phase 1b: protocol validation framework

A single demonstration of inter-rater reliability is necessary but insufficient to establish a reusable methodology. Model outputs may reflect idiosyncratic training data, vendor-specific alignment objectives, or stochastic artifacts rather than stable semantic judgments. To establish boundary conditions for a valid application, we conducted systematic validation across four dimensions.

Design Protocol. The validation manipulated four dimensions, each targeting a distinct threat to protocol validity:

- *Architecture.* Inclusion of frontier models from three independent organizations (Anthropic, OpenAI, Google) guards against vendor-specific alignment bias. Convergence across architectures provides evidence that mappings reflect semantic structure rather than training artifacts.
- *Generation.* Comparison of Phase 1a models against current-generation successors (Opus 4.5, GPT-5.2 Pro, Gemini 3 Pro Preview) assesses temporal stability. A valid protocol should produce consistent mappings across model generations.
- *Inference Mode.* Comparison of standard inference against reasoning-enhanced inference (Extended Thinking) evaluates whether deliberative processing produces systematically different judgments.
- *Temperature.* Comparison of deterministic inference (temperature = 0.0) against high-entropy inference (temperature = 1.0, three independent trials per model) assesses signal durability under stochastic variation. The latter condition is particularly important given that default user interfaces typically operate at non-zero temperature.

Phase 1b models were accessed via the OpenRouter API, which permits standardized parameter specification across providers. All API calls employed both a task-configured system prompt and the identical

user prompt from Phase 1a. The complete system prompt is reproduced in Appendix B. Table 2 presents the model configurations in Phases 1a and 1b

Validation Analyses. Eight analytical approaches assessed protocol robustness across the four design dimensions. Complete results are summarized in Table 3.

- **Inter-Rater Reliability.** Inter-Rater Reliability. Intraclass Correlation Coefficients remained excellent across all conditions (ICC range: 0.949–0.990), with cross-generational reliability maintained or improved (Phase 1a: 0.967; Phase 1b: 0.988).
- **Generalizability Theory.** Variance decomposition attributed 95.4 % to items and less than 1 % to raters ($G = 0.988$), confirming that disagreement reflects category-level judgment rather than systematic rater tendencies.
- **Equivalence Testing (TOST).** Human and AI ratings exhibited zero mean difference, with formal equivalence confirmed at $\delta = \pm 15$ % ($p = .049$).
- **Bland-Altman Analysis.** No systematic bias (mean difference = 0.0) or proportional bias ($r = -0.30$, $p = .20$) was detected; 95 % limits of agreement spanned ±18.6 percentage points.
- **Mixed-Effects Modeling.** No overall rater effects reached significance; however, GPT exhibited a category-specific Directive bias ($\beta = +17.5$, $p = .001$), part of a broader pattern of attribution bias examined in the Discussion.
- **Task Determinacy.** Duplicate trials at temperature = 0.0 produced zero inter-trial variance across all 60 ratings, confirming that the task is fully specified and admits a unique solution for each model.
- **Stochastic Robustness.** Stochastic robustness was assessed across all 27 combinations of temperature = 1.0 trials. System-level reliability remained excellent (worst-case ICC(2,1) = 0.92), and even the most divergent individual pairing maintained strong agreement (worst-case Human-AI $r = 0.87$).
- **Adversarial Weighting.** Discipline-level scores computed under five alternative weighting schemes (equal, human-only, LLM-only, 70/30, 30/70) yielded pairwise rank correlations exceeding $\rho = 0.95$. Human-only and LLM-only scoring produced $\rho = 0.983$, confirming that conclusions are invariant to rater composition.

3.2. Phase 2: derivation of AI interaction scores

This phase applied the validated consensus mapping to derive discipline-specific AI Interaction Scores, quantifying each field's orientation toward career-relevant AI usage.

The consensus allocations from Phase 1 establish the semantic correspondence between student AI interaction patterns and the five 2ACT career-relevant patterns. To derive discipline-level profiles, these allocations were applied as weights to the observed distribution of student patterns within each discipline [11]. For each discipline, the derived prevalence of each 2ACT pattern was computed as the weighted sum of observed student pattern frequencies, yielding a five-element profile representing the discipline's orientation across Directive AI, Feedback Loop AI, Validation AI, Task Iteration AI, and Learning AI.

To summarize net orientation, we computed an AI Interaction Score:

$$Score_{AI} = (P_{VA} + P_{TIA} + P_{LA}) - (P_{DA} + P_{FLA})$$

Table 2
Phase 1a and Phase 1b LLM configurations.

Dimension	Phase 1a (Baseline)	Phase 1b (Systematic Variation)
Models	Anthropic: Claude Sonnet 3.7 OpenAI: GPT-o3 Pro Google: Gemini 2.5 Pro Preview 05–06	Anthropic: Claude Opus 4.5 OpenAI: GPT-5.2 Pro Google: Gemini 3 Pro Preview
Access	Web interface	OpenRouter API
Temperature	Default (stochastic)	0.0 (× 2 trials) and 1.0 (× 3 trials)
Inference Mode	Default	Standard; Thinking ON/OFF (Opus)
Prompting	User prompt only	System prompt + User prompt

Note. Phase 1a models were accessed via provider web interfaces under default settings to emulate a standard researcher workflow. Claude Sonnet 3.7 had Extended Thinking enabled. Phase 1b models were accessed via the OpenRouter API to enable standardized parameter specification across providers. Temperature = 0.0 trials assessed task determinacy; temperature = 1.0 trials assessed stochastic robustness.

Table 3
Protocol validation framework and results.

Component	Metric	Result	Interpretation
Design Dimensions			
Architecture (cross-vendor)	Inter-model r	$\rho = 0.94$–0.99	Convergence across vendors confirms semantic structure
Generation (temporal stability)	Within-vendor r	$\rho = 0.93$–0.97	Mappings stable across model generations
Generation (temporal stability)	Cross-generational ICC	0.967 → 0.988	Reliability maintained or improved
Inference Mode (reasoning effects)	Entropy reduction	−0.61 bits	Thinking mode produces categorical judgments
Inference Mode (reasoning effects)	Human agreement	$\Delta r = +0.032$	Thinking mode aligns better with human
Temperature (stochastic robustness)	$T = 0$ vs $T = 1$ correlation	$\rho > 0.98$	Signal preserved under stochastic variation
Validation Analyses			
Inter-Rater Reliability	ICC(3,k) range (all conditions)	0.95–0.99; $M = 0.98$	Excellent reliability throughout
Generalizability Theory	Variance decomposition	Items: 95.4 %; Raters: <1 %	Disagreement is item-specific
Generalizability Theory	G coefficient	$G = 0.988$	Protocol generalizes to similar items
Equivalence Testing (TOST)	Mean difference	0.00	No systematic human-AI offset
Equivalence Testing (TOST)	Equivalence at $\delta = \pm 15$ %	$p = .049$	Formal equivalence confirmed
Bland-Altman Analysis	Systematic bias	0.0	No directional bias
Bland-Altman Analysis	95 % Limits of Agreement	[−18.6, +18.6]	Item-level variation symmetrically distributed
Bland-Altman Analysis	Proportional bias	$\rho = -0.30$ ($p = .20$)	No funneling at scale extremes
Mixed-Effects Modeling	Overall rater effects	All $p > .20$	No systematic rater severity differences
Mixed-Effects Modeling	Directive category (GPT)	$\beta = +17.5$ ($p = .001$)	Category-specific vendor bias detected
Task Determinacy	Duplicate trials ($T = 0.0$)	Zero variance	Task fully specified; unique solution
Stochastic Robustness	Worst-case ICC (2,1)	0.92	Protocol stable under maximum entropy
Stochastic Robustness	Worst-case Human-AI r	$\rho = 0.87$	Pairwise agreement preserved under entropy
Adversarial Weighting	Human-only vs LLM-only	$\rho = 0.983$	Near-perfect ranking agreement
Adversarial Weighting	All weighting schemes	$\rho > 0.95$	Conclusions invariant to rater composition

Note. Design dimensions represent the four methodological inputs varied to stress-test the protocol. Validation analyses represent the eight statistical approaches used to assess protocol robustness. Phase 1b models (Opus 4.5, GPT-5.2 Pro, Gemini 3 Pro Preview) assessed at temperature = 0.0 unless otherwise specified. Stochastic robustness assessed at temperature = 1.0 with three independent trials per model. Category-specific vendor bias (GPT Directive effect; 0.87 is the worst-case pairwise Human-GPT correlation) is examined in the Discussion.

Positive scores indicate net orientation toward augmentation-focused AI usage; negative scores indicate net orientation toward automation-focused usage. This formulation operationalizes the 2ACT framework's empirical findings: augmentation-associated patterns (Validation AI, Task Iteration AI, Learning AI) correlate with higher job zone placement and career advancement, whereas automation-associated patterns (Directive AI, Feedback Loop AI) correlate with lower job zone placement and routine task susceptibility.

The derived AI Interaction Scores are presented in Table 4 (Section 4).

3.3. Phase 3: external indicator analysis

This phase examined associations between derived AI Interaction Scores and real-world indicators of labor market outcomes and demographic composition.

3.3.1. Data sources

For each of the nine academic disciplines, data were collected on two external indicators:

Median Annual Earnings. Median annual earnings for full-time, year-round workers aged 25 to 64 holding a bachelor's degree in the relevant field, sourced from the U.S. Department of Education, National Center for Education Statistics (NCES), Digest of Education Statistics 2022, Table 505.06 [54].

Gender Composition. Percentage of bachelor's degrees conferred to female students by field of study in the 2021–22 academic year, sourced from NCES Digest of Education Statistics 2022, Table 318.30 [53].

3.3.2. Data harmonization

A structural mismatch between data sources required harmonization. Handa et al. [11] report Natural Sciences and Mathematics as a single aggregated category, whereas NCES provides disaggregated data for constituent fields. To establish accurate demographic and economic profiles for this combined category, we calculated weighted averages. For percent female graduates, weights were derived from bachelor's degree conferral counts for Natural Sciences and Mathematics from NCES Table 318.30 [53]. For median annual earnings, weights were derived from the number of 25- to 64-year-old bachelor's degree holders in each field from NCES Table 505.06 [54]. The use of bachelor's degree data as the primary benchmark aligns with the comparative metric employed by Handa et al. [11]. We acknowledge that the source data comprise an unsegmented mix of undergraduate and graduate students; this constraint is addressed in the limitations.

3.3.3. Analytical approach

Given the preliminary nature of this phase and the characteristics of population-level analysis with nine units, we employed nonparametric methods that make no assumptions about functional form or distributional properties.

Visualization. Scatter plots were constructed to examine relationships between AI Interaction Score and Median Annual Earnings, AI Interaction Score and Percent Female Graduates, and Percent Female Graduates and Median Annual Earnings. These visualizations permit inspection of functional form, identification of potential outliers, and assessment of linearity assumptions.

Correlation Analysis. Spearman's rank correlation coefficients (ρ) were computed for each bivariate relationship. Spearman's ρ was selected over Pearson's r because it makes no assumptions about functional form, is robust to outliers, and supports ordinal interpretation. An alpha level of 0.05 was used for significance assessment. All analyses were conducted using Python 3.13 with the SciPy statistical library.

The intent of these analyses is to identify structural patterns and potential covariations that warrant further investigation, not to establish causality. Findings should be interpreted as hypothesis-generating rather than confirmatory.

3.3.4. Robustness assessment

To ensure stability and reliability of observed associations, four complementary robustness procedures were applied to each correlation:

Bootstrap Analysis. 95 % confidence intervals were generated using 10,000 bootstrap resamples with replacement, providing nonparametric uncertainty estimates independent of distributional assumptions.

Rater Independence Sensitivity. To test whether associations with external indicators reflect structural relationships rather than consensus artifacts, we examined correlations at both rater and weighting levels. At the rater level, each rater's independent allocations were propagated through the AI Interaction Score formula, and correlations with external indicators were computed separately for each of the four raters. All four raters produced significant correlations with both outcome variables in the same direction (Percent Female Graduates: all $\rho < -0.66$, $p < .05$; Median Annual Earnings: all $\rho > 0.71$, $p < .05$). This four-way independent replication provides strong evidence that observed associations reflect structural correspondence between AI usage patterns and external indicators rather than idiosyncratic features of any single rater's judgment. At the consensus level, correlations were recomputed using AI Interaction Scores derived under five alternative weighting schemes (equal weights, human-only, LLM-only, 70/30 human/LLM, 30/70 human/LLM). Associations with both external indicators remained significant across all weighting configurations.

Leave-One-Out Sensitivity. Each correlation was recomputed nine times, excluding one discipline per iteration. Stable correlations across all leave-one-out conditions indicate that findings are not artifacts of individual influential observations.

Disaggregation Sensitivity. Correlations involving the aggregated Natural Sciences and Mathematics category were recomputed using disaggregated NCES data, treating Natural Sciences and Mathematics as separate observations. Concordance between aggregated and disaggregated results supports the validity of the harmonization procedure.

Table 4
Consolidated disciplinary profiles: AI interaction patterns, economic outcomes, and demographics.

Discipline	DA	FLA	VA	TIA	LA	AI	Earnings	% Fem.
Natural Sci & Math	26.15	4.88	8.41	28.08	32.65	38.11	$88,870	46.52 %
Computer Science	26.62	5.10	9.13	30.33	28.98	36.72	$105,890	22.62 %
Engineering	26.88	5.10	9.76	31.77	26.67	36.23	$114,510	25.28 %
Psychology	30.26	4.95	9.39	29.31	26.22	29.71	$71,860	80.11 %
Social Sci & History	31.21	5.07	9.58	29.79	24.60	27.69	$88,085	53.03 %
Humanities	32.47	4.95	9.22	28.15	25.47	25.41	$71,090	64.86 %
Business	32.54	4.97	9.15	28.03	25.44	25.11	$88,560	46.48 %
Health Professions	33.36	4.95	9.03	27.39	25.32	23.43	$83,170	84.89 %
Education	34.05	5.15	10.14	30.46	20.36	21.75	$62,420	83.19 %

Note. DA = Directive AI; FLA = Feedback Loop AI; VA = Validation AI; TIA = Task Iteration AI; LA = Learning AI. Earnings = median annual earnings for bachelor's degree holders aged 25–64 (NCES, 2023b). % Female = percentage of bachelor's degrees conferred to women (NCES, 2023a). Rows sorted by AI Score (descending).

3.4. Ethical considerations

This study utilized publicly available and aggregated data sources throughout. The student AI interaction data from Handa et al. [11] were anonymized and filtered by the original authors prior to public release. The demographic and earnings data from NCES are published government statistics aggregated at the discipline level. No primary data were collected from human subjects, and no individually identifiable information was accessed at any stage of the research.

The use of large language models as raters involved no human subject considerations, as LLMs are computational tools rather than research participants. Throughout the rating process, the authors maintained full human oversight and control. All model outputs were reviewed for coherence and appropriateness, and the authors take full responsibility for the content and interpretation of the findings.

3.5. Data and code availability

To support transparency and replication, the following materials are provided as supplementary files with this article: (a) the complete prompts used for LLM rating (Appendix A user prompt and Appendix B system prompt), (b) the raw ratings from all human and LLM raters across both protocol phases, (c) the derived discipline-level AI Interaction Scores, (d) the compiled dataset with AI Interaction Scores and external indicators, and (e) the Python analysis script used to produce all findings, figures, and tables reported in this study. A data dictionary defining all variables accompanies the supplementary materials. The primary data on student AI interactions are publicly available from Handa et al. [11], and the demographic and earnings data are available from the National Center for Education Statistics (NCES, 2023a; 2023b). The derived data and analysis script used to generate all reported findings are also available at OSF DOI [to be updated upon acceptance].

4. Results

Results characterize associations across the complete NCES field of study taxonomy (nine disciplines). Although derived from 574,740 student conversations, analyses operate at the discipline level. All associations are descriptive; no causal inference is claimed. Because these nine disciplines constitute the complete taxonomy, reported statistics describe observed population relationships rather than estimated parameters for generalization. To control for the false discovery rate across the three primary correlational tests (Gender-AI Score, Earnings-AI Score, Gender-Earnings), we applied the Benjamini-Hochberg procedure to the calculated p-values.

4.1. The gender equity disparity: AI interaction scores negatively correlate with female representation

A statistically significant negative correlation emerged between AI Interaction Scores and percentage of female graduates ($\rho = -0.68$, $p = .042$, $p_{adj} = 0.042$). Using the conventions proposed by Cohen [52] for interpreting effect sizes, this represents a large effect ($|\rho| > 0.50$). The stability of this finding was confirmed via bootstrap analysis, which yielded a 95 % confidence interval of [−0.95, −0.01]. A leave-one-out sensitivity analysis also demonstrated that the correlation was not driven by any single discipline.

This relationship manifests starkly across disciplines. Education, with 83.19 % female graduates, recorded the lowest AI Interaction Score (21.75), while the gender-balanced Natural Sciences & Mathematics, with 46.52 % female graduates, achieved the highest score (38.11). Computer Science, with only 22.62 % female graduates, also scored high (36.72). The pattern persists across the disciplinary spectrum, as visualized in Fig. 2.

4.2. Career implications: AI scores correlation with disciplinary earnings

Complementing the gender finding, AI Interaction Scores show a statistically significant, strong positive correlation with median annual earnings ($\rho = 0.72$, $p = .030$, $p_{adj} = 0.042$). This large effect size, confirmed with a bootstrapped 95 % confidence interval of [0.13, 0.95], indicates that disciplines whose aggregate AI usage patterns favor augmentation-focused interactions are also associated with substantially higher median earnings among degree holders in those fields.

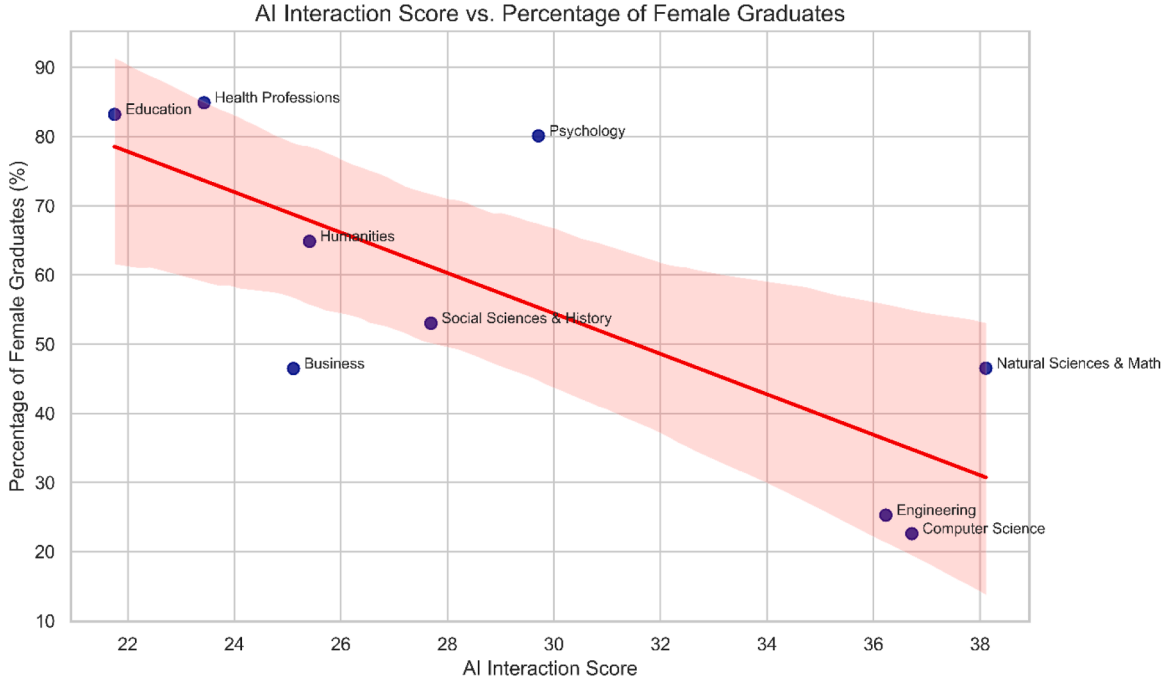

Fig. 2. Scatter Plot of AI Interaction Score vs. Percentage Female Graduates for $N = 9$ Majors
Note. The plot shows the relationship for the $N = 9$ academic majors, illustrating the strong negative correlation ($\rho = -0.68$). The solid red line is the linear regression fit, and the shaded area represents the 95 % confidence interval for the regression estimate.

The earnings differential is substantial: disciplines in the highest tertile of AI Interaction Scores (Natural Sciences & Math, Computer Science, Engineering) report median earnings of $88,870 to $114,510, while those in the lowest tertile (Business, Health Professions, Education) range from $62,420 to $88,560. Table 4 juxtaposes each discipline's AI interaction profile, score, salary, and gender mix, making the economic implications clear. The gap between the highest-earning discipline, Engineering ($114,510), and the lowest-earning, Education ($62,420), is $52,090, a disparity visualized by the upward trend in Fig. 3. The scatter plot shows how disciplines align along this gradient, with STEM fields clustering in the high-score, high-earnings quadrant while fields like Education and Humanities occupy the opposite corner.

4.3. The intersection: female-dominated disciplines face dual disadvantage

A third significant correlation reveals the compounding nature of these disparities: the percentage of female graduates negatively correlates with median annual earnings ($\rho = -0.85$, $p = .004$, $p_{adj} = 0.011$). This large effect, supported by a bootstrapped 95 % confidence interval of [−1.00, −0.37], suggests that female-dominated disciplines face both lower earning potential and, as previously shown, map to AI usage patterns less conducive to career advancement [14]. This represents a dual disadvantage with profound implications for perpetuating gender-based economic disparities in an AI-driven economy.

4.4. Mapping student behaviors to career competencies reveals mixed potential

The systematic translation of student AI patterns to career-relevant competencies yielded nuanced insights. Through multi-rater conceptual mapping that achieved excellent reliability (ICC values 0.949–0.984, all $p < .0001$), clear distinctions emerged in how different student interaction modes align with career outcomes. DOC mapped overwhelmingly to Directive AI (76.3 %), the pattern most strongly associated with lower occupational placement. Conversely, COC aligned primarily with augmentation-focused patterns, namely Task Iteration AI (62.5 %) and Validation AI (18.8 %). DPS showed a dual character, splitting almost evenly between the career-advancing Learning AI (48.8 %) and the automation-focused Directive AI (46.3 %). Finally, CPS also strongly favored career-advancing patterns (Table 1).

4.5. Two-Thirds of student AI interactions show augmentation potential but automation persists

Weighted analysis across disciplines reveals cautious optimism tempered by observed blunting tendencies. Approximately 66.5 % of student AI interactions align with augmentation-focused patterns (Task Iteration: 29.5 %, Learning: 27.9 %, Validation: 9.1 %), suggesting latent potential for career-advancing engagement. However, 33.6 % remain automation-focused, with Directive AI comprising 28.6 % and Feedback Loop AI comprising 5.0 % of all interactions. A substantial proportion of students engage with AI in ways associated with skill displacement rather than enhancement (Table 4; Fig. 4). Notably, Validation AI comprises only 9.1 % of interactions, indicating that students rarely leverage AI's capacity to verify and improve their own work, representing another missed opportunity for augmentation-focused learning.

5. Discussion

This study's findings should not be discounted as merely observational; they highlight emerging challenges thrust upon higher education by the proliferation of AI technologies. The following interpretation operates at the discipline level; no claims are made about individual student behavior or outcomes. By systematically mapping disciplinary AI interaction patterns to empirically validated career competencies, we observe evidence of an emerging, yet largely invisible, systemic algorithmic gender gap in higher education. The data suggest an architecture of disadvantage: disciplines with higher female representation are associated with AI usage patterns consistent with skill automation rather than augmentation. These same interaction patterns are strongly correlated with lower graduate earnings, suggesting existing disciplinary differences in AI engagement may amplify rather than ameliorate gender-based economic disparities. While two-thirds of student

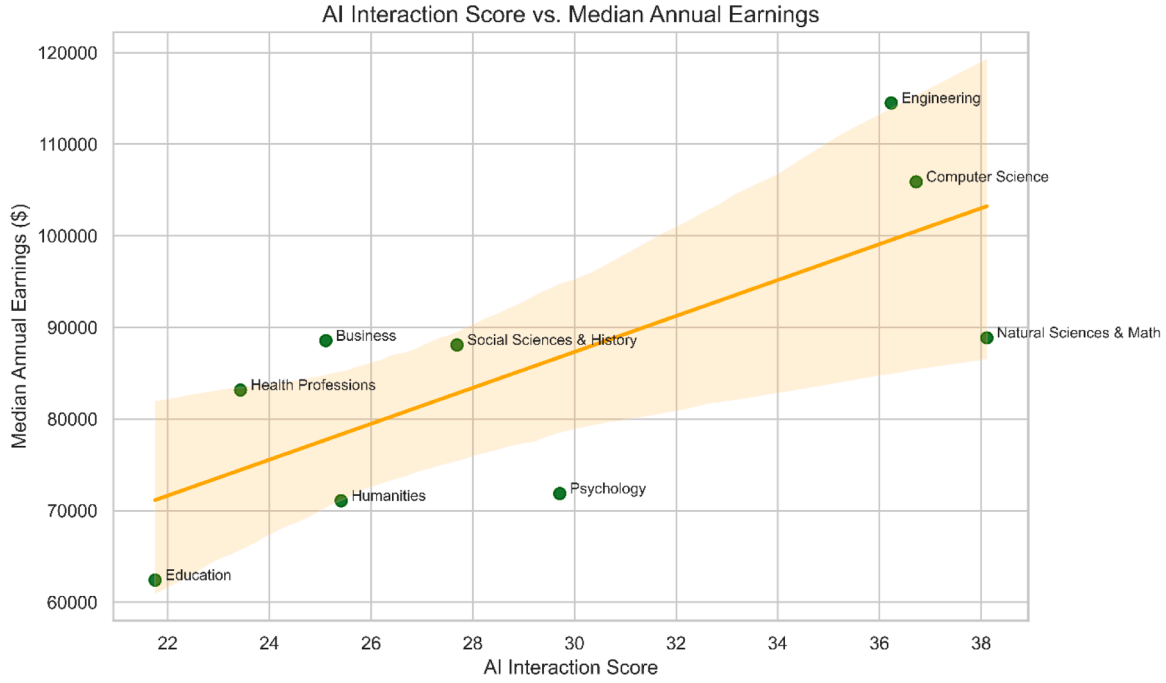

Fig. 3. Scatter Plot of AI Interaction Score vs. Median Annual Earnings
Note. The plot shows the relationship between academic disciplines' AI scores and median annual earnings, illustrating the strong positive correlation ($\rho = 0.72$). The solid orange line is the linear regression fit, and the shaded area represents the 95 % confidence interval for the regression estimate.

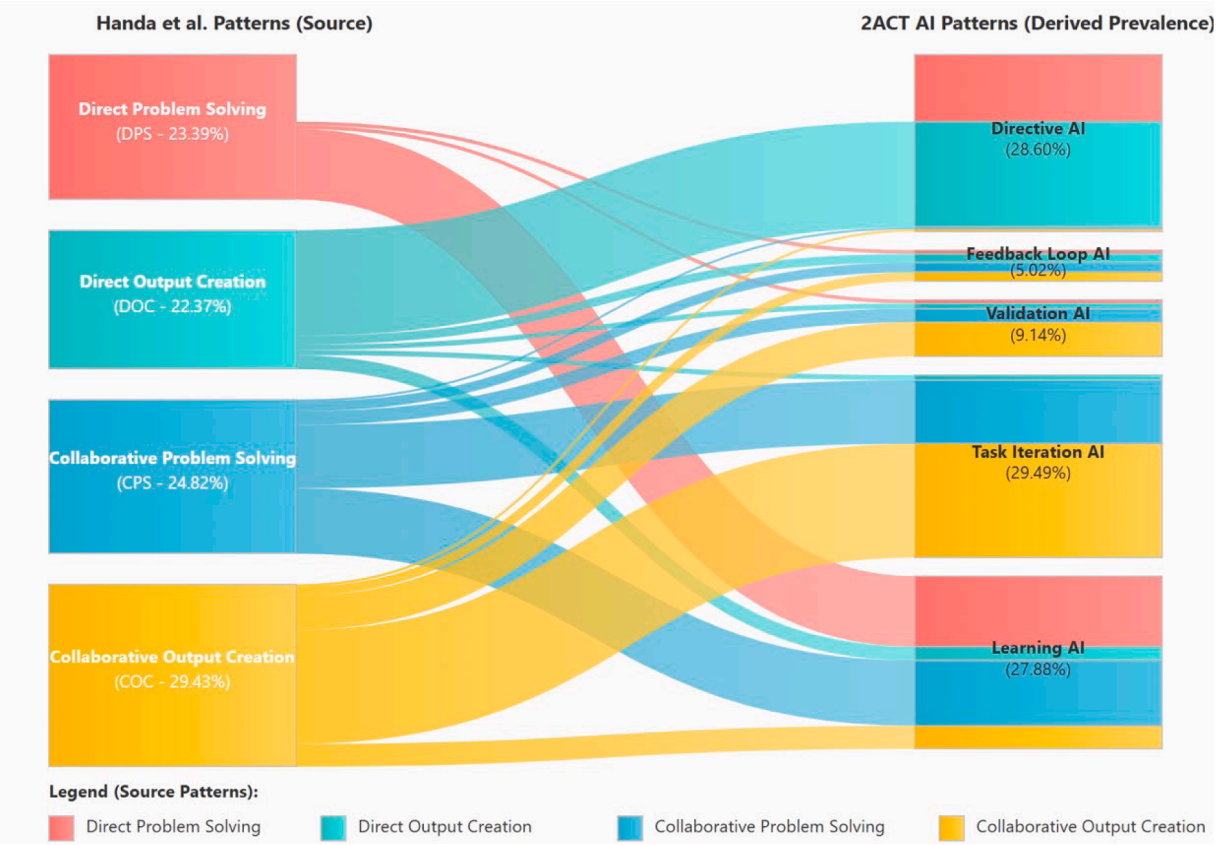

Fig. 4. Sankey Diagram Mapping Student AI Usage to Career-Relevant Patterns
Note. This flow diagram illustrates the conceptual mapping from the four student AI interaction patterns observed by Handa et al. [11] (left column) to the five career-relevant 2ACT patterns (right column). The width of each flow is proportional to the percentage of a student's pattern that was allocated to a corresponding career pattern during the multi-rater analysis. The percentages on the right indicate the final derived prevalence for each career-relevant pattern across all student interactions.

interactions show latent potential for career advancement, the persistence of automation-focused disciplinary AI patterns, particularly in female-dominated fields, represents an important though addressable challenge.

The validation methodology revealed an unexpected dimension of this bias. All three AI models exhibited attribution bias: a systematic tendency to rate AI-mediated interactions as more controlling and less pedagogically oriented than the human researcher rated them. This inverts the commonly hypothesized Benevolence Bias, wherein AI models might be expected to rate AI favorably. Instead, AI models appear to perceive AI (themselves) as more determinative of interaction outcomes than humans perceive. If LLM-based analysis tools are deployed to evaluate educational AI integration, this bias could systematically overestimate the automation-focused character of student interactions, potentially distorting institutional assessments and policy responses. The multi-rater consensus protocol developed here mitigates this risk, but single-model evaluations warrant caution.

5.1. Theoretical implications

These findings extend Vygotsky's [32] Zone of Proximal Development (ZPD) into the AI era.

5.1.1. A bifurcated ZPD\

In disciplines like Engineering and Computer Science, AI appears to operate as a cognitively scaffolded collaborative partner, extending the ZPD by enabling students to tackle more complex iterative and validation-oriented tasks (high Task Iteration and Validation AI). Conversely, in disciplines like Education and Health Professions, AI may more often be leveraged as a substitute for cognitive effort (high Directive AI), effectively collapsing the ZPD and promoting dependency over capability. We contend the problem is not the tool but the sociotechnical context that normalizes its use for cognitive offloading rather than cognitive expansion.

5.1.2. A new vector for systemic bias

This research advances critical AI scholarship [16,17] by identifying a novel mechanism through which inequality can be perpetuated. Beyond biased datasets or algorithmic outputs, our findings suggest inequity may be encoded in patterns of human-AI interaction that disciplines pedagogically promote or tacitly permit. If so, this is a subtler, more deeply embedded form of algorithmic bias that operates not at the level of code but at the level of classroom culture and assessment, making it uniquely an educational challenge.

5.1.3. Extending the augmentation/automation duality

The 2ACT framework established the augmentation/automation distinction and the association with outcomes in occupational contexts. This study suggests that disciplines exhibiting augmentation-aligned educational patterns are also associated with higher earnings, indicating this distinction operates across domains with implications for educational design and lifelong earning potential.

5.2. The gender equity disparity: A compounding disadvantage

The strong negative correlation between female representation and AI interaction scores ($\rho = -0.68$) is the study's most consequential finding. When cross-referenced with the strong negative correlation

between female representation and earnings ($\rho = -0.85$), a dual disadvantage becomes evident. Female-dominated fields like Education ($62,420 median earnings) and Health Professions ($83,170) already face substantial wage gaps compared to male-dominated fields like Engineering ($114,510). Our analysis reveals that AI educational integration, in its current state, is associated with lower pay and lower occupational mobility [14], potentially positioning AI to reinforce existing labor stratification rather than disrupt it.

These findings align with external indicators that suggest the gap may be widening. Freeman [3] reports that male students use AI to "improve AI skills" at 64 % higher rates than female students. This disparity appears structurally mediated: female students are 38 % more likely to report that their institution bans or discourages AI and 31 % more likely to express concern about being accused of cheating. Institutional encouragement rates favor male students by 45 %. Freeman's [3] policy note explains, "The digital divide we identified in 2024 appears to have widened" (p. 1). These conditions create what might be termed a compliance asymmetry: students in environments with higher encouragement are freer to explore iterative augmentation patterns, while students in more restrictive environments are channeled toward efficient, transactional modes that minimize detection risk.

This is a systemic challenge, not an individual deficiency. It comes as no surprise that female students facing lower directed use of AI and higher blanket bans express 39 % lower confidence in AI [46]. The consequence is seemingly a reinforcing cycle: restrictive policies encourage covert, transactional AI use, which prevents development of the overt augmentation skills the labor market increasingly demands, which in turn leaves students less prepared to advocate for constructive AI integration. Thus, students in fields critical to social infrastructure are immersed in environments that normalize covert task completion while discouraging the overt AI competencies demanded by labor markets.

The contrast between potential and practice underscores our central argument: the pedagogical framework surrounding AI may determine whether it becomes a tool for stratification or equity.

5.3. Methodological implications: toward replicable LLM-assisted research

Beyond substantive findings, this study contributes a validated protocol for LLM-assisted qualitative coding that addresses growing concerns about reproducibility and validity in AI-augmented research. The rating protocol treats LLM variance as an experimental variable, systematically varying architecture, generation, inference mode, and temperature to isolate the conditions under which human and AI raters converge or diverge. The validation protocol moves beyond single-metric reliability to characterize the structure of disagreement through eight complementary analyses, including Generalizability Theory, equivalence testing, Bland–Altman analysis, mixed-effects modeling, and adversarial weighting.

5.3.1. The protocol as contribution

The multi-rater framework developed herein that integrates human expertise with architecturally diverse LLMs under controlled inference conditions may serve as a template for researchers seeking to combine LLM efficiency with psychometric rigor. The validation phase demonstrates that such protocols can achieve excellent reliability (ICC(3,k) > 0.95 across all conditions), confirmed equivalence between human and AI ratings (TOST $p = .049$), and robustness to stochastic variation (worst-case ICC(2,1) = 0.92 across 27 trial combinations). The protocol produces identical substantive conclusions regardless of how raters are weighted: human-only and LLM-only scoring yielded $\rho = 0.983$ in discipline rankings. This invariance to rater composition provides strong evidence for construct validity.

5.3.2. The calibration paradox

A counterintuitive finding warrants attention from researchers employing LLMs. Across model generations, correlation and reliability improved while category-level calibration worsened. Current-generation models exhibited higher inter-rater agreement than their predecessors, yet also showed larger deviations from human baseline ratings. GPT's Directive bias tripled from +5 % to +17.5 % across generations. This dissociation, wherein models become more internally consistent but more extreme in their judgments, suggests that alignment training may optimize for decisiveness at the cost of nuanced probability distributions. The implication is that newer models are not necessarily more valid for rating tasks; cross-generational validation should become standard practice.

5.3.3. Attribution bias

All AI models rated AI-mediated interactions as more directive and less learning-oriented than those of the human researcher. This systematic bias manifests as capability self-attribution: AI models perceive AI as more determinative of interaction outcomes than humans perceive it to be. The bias persists across architectures and generations, suggesting it may reflect structural properties of how language models represent agency rather than idiosyncratic training artifacts. For applied research, the implication is clear: single-model coding will systematically overestimate automation-focused usage. Multi-rater consensus with human anchoring remains essential.

5.3.4. Task determinacy as validity evidence

The finding that duplicate deterministic trials produced zero variance across all 60 ratings provides a novel form of validity evidence. In open-ended generative tasks, temperature = 0.0 would still produce inter-trial variance as multiple token sequences compete at the initial sampling step. The absence of such variance confirms that the conceptual mapping task, as specified, admits a unique optimal solution for each model. This task determinacy indicates that observed ratings reflect stable semantic judgments rather than stochastic artifacts, and that the prompt specification was sufficient to constrain the output space. We recommend that future LLM-assisted research protocols include determinacy verification as standard practice.

5.3.5. Limitations of LLM agreement

The fundamental question of whether LLM agreement reflects genuine semantic understanding or sophisticated pattern matching cannot be definitively resolved. The structural correspondence between educational and occupational frameworks may be recoverable through pattern matching alone, which would not invalidate the mapping but would limit claims about the depth of conceptual analysis involved. The convergence of human and LLM ratings provides pragmatic validation; the protocol produces reliable, replicable results without requiring resolution of deeper questions about machine understanding.

6. Pedagogical recommendations: from passive consumption to active collaboration

While longitudinal confirmation remains necessary, the strength and consistency of observed associations warrant precautionary pedagogical action. Established correlations between AI interaction patterns and labor market outcomes [14], combined with the disparities documented here, suggest that inaction carries its own risks. We propose five evidence-informed recommendations:

6.1. Reconsider AI prohibitions

With 92 % of undergraduates reporting AI usage, 88 % employing generative AI for assessments [3], and 74 % failing to disclose usage despite mandatory requirements [47], prohibition may function as de facto mandated concealment. Indeed, Ha et al. [46] report that 65 % of students admit to cheating with AI in school. Prohibition seemingly does not reduce AI engagement, though it ensures engagement occurs

without guidance, scaffolding, or skill development. The question is not whether students use AI but whether institutions shape that use.

6.2. Critically evaluate the integrity–authenticity dichotomy

Prevailing integrity frameworks treat AI assistance as inherently antithetical to authentic learning [48]. However, empirical evidence contradicts the assumption that AI use necessarily undermines learning. Preliminary findings suggest that engagement and motivation are improved with AI integration [49], with meta-analytic evidence indicating improved academic performance [50]. While some may find the calculator analog informative, where the tool redirected cognitive resources toward higher-order tasks, we contend that AI is part and parcel of the modern, professional workplace. AI competency development should be an element of integrity-based learning frameworks to ensure market alignment. The pedagogical question is not whether AI undermines learning but whether institutions design for augmentation. Integrity must evolve from prohibition to proficiency.

6.3. Consider process in assessment

Our findings reveal that students engage AI across both augmentation and automation modes; both have legitimate applications requiring distinct competencies. The data show 66.5 % of interactions carry augmentation potential while 33.6 % remain automation-focused. The issue is not eliminating automation but ensuring intentional deployment effectively. Delegating thinking without justification presents documented risks, most notably capability impairment [35]. Conversely, appropriately scaffolded AI integration enables students to maintain cognitive engagement while leveraging AI capabilities [50]. Assignment design should make both modes visible, teaching students to recognize when delegation is appropriate and when iteration adds value. As long as assessment focuses solely on the final deliverable, students may be incentivized to use AI for automation and cognitive offloading. Transforming assessment to value the process of collaboration serves to cultivate career-advancing skills. This could involve grading students' prompt chains, their analysis of AI-generated feedback, or a reflective essay on how they strategically leveraged AI as a cognitive partner.

6.4. Elevate critical thinking

Critical thinking without the critical is opinion. Yet, AI invites a peculiar abdication: the outsourcing of judgment to systems that present probabilistic outputs as authoritative conclusions. We contend that critical evaluation is not merely preserved but elevated in the AI era. The data support this urgency. Our findings indicate only 9.1 % of student interactions involve Validation AI, suggesting students rarely leverage AI to critique and improve their own work. Broader evidence confirms the pattern: 66 % of users accept AI output without evaluating accuracy [35], and only 23 % of Americans demonstrate high AI knowledge [46]. Students are engaging tools they do not understand and accepting outputs they do not verify.

Error-detection and claim-verification competencies must be explicitly taught and independently assessed. Students require training to identify errors, evaluate claims against evidence, and recognize the boundaries of AI reliability. These skills do not emerge from usage alone; they emerge from intentional pedagogical design.

6.5. Evaluate and execute

Any policy is defensible absent a defined destination. Only 20 % of institutions have published AI policies, and just 14 % have reviewed curriculum to align with the rise of AI in the workplace [6]. Yet, 90 % of higher education professionals use AI tools despite over half having uncertainty about effective application [7]. The consequence: the majority of recent graduates report their programs failed to prepare them for generative AI, with three-quarters requiring additional training in current roles [51]. Institutions cannot address what they do not see. Surfacing AI interaction patterns transforms speculation into strategy; without visibility, resource allocation is guesswork. The only durable strategy is the commitment to continuous measurement, intervention, and refinement.

7. Limitations and future research

This study, despite establishing significant relationships, has inherent limitations that both contextualize the findings and pave the way for future research.

7.1. Analytic scope

The nine disciplines examined constitute the complete NCES broad-field taxonomy, representing a population enumeration rather than a sample. However, this aggregation level necessarily obscures within-discipline heterogeneity. Future research using institution-level or individual-level data could examine variation in AI usage patterns within fields, including potential gender differences in AI engagement among students in the same discipline.

7.2. Data source constraints

A structural limitation exists within the foundational Handa et al. [11] data. Derived from general university accounts, the source population likely contains an unsegmented mix of undergraduate and graduate students who may exhibit systematically different interaction patterns. While this study aligned its demographic and economic benchmarks with the bachelor's degree level used by Handa et al. for methodological consistency, this introduces the risk of ecological fallacy. Graduate students engaged in research may naturally perform more augmentation-focused tasks, potentially inflating AI Interaction Scores for disciplines with high graduate enrollment. Future research using AI interaction data explicitly segmented by academic level is needed to disentangle these effects.

7.3. Protocol validation boundaries

The systematic validation study confirmed excellent reliability across model architectures, generations, and inference conditions. However, several boundary conditions warrant acknowledgment. First, the protocol included only one human rater; while leave-one-out analysis confirmed the human was not an indispensable anchor, a multi-human panel would strengthen validity claims. Second, the validation identified systematic vendor biases: all AI models exhibited attribution bias (rating AI as more directive than the human researcher), with GPT showing the largest effect ($\beta = +17.5$, $p = .001$). While multi-rater consensus dilutes these biases, studies employing single-model coding should interpret results with caution. Third, a Calibration Paradox was observed across model generations, wherein correlation improved while category-level bias increased, suggesting that ongoing validation will be necessary as frontier models continue to evolve. The rater and validation frameworks reported herein provide a template for such continued assessment.

7.4. LLM methodology

The use of LLMs as raters in conceptual analysis remains an emerging methodology. Although task determinacy was confirmed (zero inter-trial variance at temperature = 0.0) and cross-generational stability was demonstrated (within-vendor $r > 0.92$), the fundamental question of whether LLM agreement reflects genuine semantic understanding or sophisticated pattern matching cannot be definitively resolved. The structural correspondence between educational and occupational

frameworks may be recoverable through pattern matching alone, which would not invalidate the mapping but would limit claims about the depth of conceptual analysis involved. We recommend that future research continue to investigate boundary conditions for LLM-assisted qualitative coding across diverse task types.

7.5. Measurement uncertainty

While the nine disciplines constitute the complete NCES taxonomy, the AI Interaction Scores assigned to each are derived quantities mediated by rater judgment rather than directly observed population parameters. Inter-rater reliability was excellent (ICC $>$ 0.94), and adversarial weighting confirmed invariance to rater composition, but item-level variation persists. Reported correlations should be interpreted with this measurement-level uncertainty in mind.

7.6. Future directions

Longitudinal research tracking graduates from this cohort into the workforce would validate whether AI Interaction Scores predict career trajectory and earnings as theorized. Additionally, the pedagogical implications outlined in the Discussion should be evaluated through experimental and quasi-experimental studies measuring the efficacy of interventions designed to shift student AI usage toward augmentation-focused patterns. Further, future studies may conceptualize automation and augmentation as independent portfolio dimensions rather than opposing poles, modeling optimal ratios across task contexts.

8. Conclusion

Despite these limitations, this study provides the first empirical bridge between how students actually use AI and the specific interaction patterns associated with economic advancement. The correlations are too strong and the potential social implications too profound to ignore. Higher education stands at a defining crossroads. It can continue the current trajectory and risk inadvertently contributing to a new algorithmic stratification system, or it can proactively intervene to ensure that AI becomes a tool for equitable advancement for all students.

CRediT authorship contribution statement

Drake Mullens: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Stella Shen:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.caeo.2026.100350.

Appendix A: Full Prompt for Conceptual Mapping Task

The following prompt was provided to the primary researcher and the three Large Language Model raters to perform the conceptual mapping of Handa et al.'s [11] student interaction patterns to the 2ACT framework [14].

FRAMEWORK 1: Handa et al. [11] **identified four student AI interaction patterns:**

1. Direct Problem Solving (DPS): Students seek direct solutions or explanations for specific questions with minimal iteration.
2. Direct Output Creation (DOC): Students request AI to generate longer outputs (e.g., essays, code, etc.) with minimal iterative engagement.
3. Collaborative Problem Solving (CPS): Students engage in dialogue with AI to understand concepts or solve problems iteratively.
4. Collaborative Output Creation (COC): Students and AI co-create or iteratively refine longer outputs through multiple exchanges.

FRAMEWORK 2: Mullens & Shen [14] identified five primary career-relevant AI interaction patterns in their 2ACT framework**:**

1. Directive AI: AI systems perform tasks with minimal human involvement (associated with lower job zones).
2. Feedback Loop AI: AI receives environmental feedback to complete tasks (associated with lower job zones).
3. Validation AI: AI validates or verifies human outputs (conditionally associated with higher job zones when paired with scientific knowledge).
4. Task Iteration AI: Humans and AI collaboratively refine outputs through multiple exchanges (associated with higher job zones).
5. Learning AI: AI provides knowledge, explanations, or training to humans (associated with higher job zones).

NOTE: The 2ACT framework also includes "Thinking Fraction" as a cross-cutting dimension measuring depth of cognitive engagement across all patterns. As this is not a separate pattern category but rather a dimension that applies across patterns, please exclude it from your primary mapping.

TASK: For each student pattern (Framework 1), assign percentage allocations to the relevant 2ACT patterns (Framework 2) that best represent their conceptual alignment. The percentages for each student pattern should sum to 100 %.

Please provide:

1. A percentage-based mapping for each student pattern
2. A brief justification for each mapping decision (2–3 sentences)
3. Confidence level in your mapping (high/medium/low)

Respond in a structured, tabular format for ease of analysis.

Appendix B: System Prompt for Protocol Validation (Phase 1b)

The following system prompt was used for all API-based inference trials during the Phase 1b validation protocol

You are an expert researcher in labor economics and educational pedagogy. You are tasked with performing a conceptual mapping between two theoretical frameworks. Your goal is to identify conceptual alignment based strictly on the provided definitions. You must provide your output in a structured markdown table. Be deterministic and objective.

References

[1] Digital Education Council. (2024). *What students want: Key results from DEC Global AI Student Survey 2024*. https://www.digitaleducationcouncil.com/post/what-students-want-key-results-from-dec-global-ai-student-survey-2024.
[2] Chegg. (2025, January 28). Chegg Global Student Survey 2025: 80% of undergraduates worldwide have used GenAI to support their studies – but accuracy a top concern [Press release]. https://investor.chegg.com/Press-Releases/press-release-details/2025/Chegg-Global-Student-Survey-2025-80-of-Undergraduates-Worldwide-Have-Used-GenAI-to-Support-their-Studies–But-Accuracy-a-Top-Concern/default.aspx.
[3] Freeman J. Student generative AI survey 2025. Higher education policy institute; 2025. https://www.hepi.ac.uk/2025/02/26/hepi-kortext-ai-survey-shows-explosive-increase-in-the-use-of-generative-ai-tools-by-students/.

[4] World Economic Forum. (2025). *The future of jobs report 2025*. https://www.weforum.org/publications/the-future-of-jobs-report-2025/.
[5] Chiu TKF. AI literacy and competency: definitions, frameworks, development and future research directions. Interact Learn Environ 2025;33(5):3225–9. https://doi.org/10.1080/10494820.2025.2514372.
[6] Inside Higher editor (2024). *2024 survey of college and university chief academic officers*. Inside Higher Ed & Hanover Research. https://www.insidehighered.com/sites/default/files/2024-05/Slides%202024%20Survey%20of%20College%20and%20University%20Chief%20Academic%20Officers.pdf.
[7] UNESCO. (2025). Global Survey AI high education. https://www.unesco.org/en/articles/unesco-survey-two-thirds-higher-education-institutions-have-or-are-developing-guidance-ai-use.
[8] Acemoglu D, Restrepo P. Robots and jobs: evidence from US labor markets. J Polit Econ 2020;128(6):2188–244. https://doi.org/10.1086/705716.
[9] Autor DH. Why are there still so many jobs? The history and future of workplace automation. J Econ Perspect 2015;29(3):3–30. https://doi.org/10.1257/jep.29.3.3.
[10] Brynjolfsson E, McAfee A. The second machine age: work, progress, and prosperity in a time of brilliant technologies. W. W. Norton & Company; 2014.
[11] Handa K, Bent D, Tamkin A, McCain M, Durmus E, Stern M, Schiraldi M, Huang S, Ritchie S, Syverud S, Jagadish K, Vo M, Bell M, Ganguli D. Anthropic education report: how university students use Claude. Anthropic; 2025. https://www.anthropic.com/news/anthropic-education-report-how-university-students-use-claude.
[12] Hirschi A. The fourth industrial revolution: issues and implications for career research and practice. Career Dev Q 2018;66(3):192–204. https://doi.org/10.1002/cdq.12142.
[13] World Economic Forum. (2023). *The future of jobs report 2023* (Insight Report). https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf.
[14] Mullens, D., & Shen, S. (2025). *2ACT: aI-accentuated career transitions via skill bridges* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2505.07914.
[15] Selwyn N. Should robots replace teachers? AI and the future of education. John Wiley & Sons; 2019.
[16] D'Ignazio C, Klein LF, Klein LF. Introduction: why data science needs feminism. In: D'Ignazio C, editor. Data feminism. The MIT Press; 2020. p. 1–20. https://data-feminism.mitpress.mit.edu/pub/frfa9szd/release/5.
[17] Noble SU. Algorithms of oppression: how search engines reinforce racism. New York University Press; 2018.
[18] Gašević D, Siemens G, Sadiq S. Empowering learners for the age of artificial intelligence. Comput Educ: Artif Intell 2023;4:100130. https://doi.org/10.1016/j.caeai.2023.100130. article.
[19] Lodge JM, de Barba P, Broadbent J. Learning with generative artificial intelligence within a network of co-regulation. J Univ Teach Learn Pract 2023;20(7):02. https://doi.org/10.53761/1.20.7.02.
[20] Salomon G. Cognitive effects with and of computer technology. Commun Res 1990;17(1):26–44. https://doi.org/10.1177/009365090017001002.
[21] Sparrow B, Liu J, Wegner DM. Google effects on memory: cognitive consequences of having information at our fingertips. Science 2011;333(6043):776–8. https://doi.org/10.1126/science.1207745.
[22] Williamson B, Eynon R. Historical threads, missing links, and future directions in AI in education. Learn Media Technol 2020;45(3):223–35. https://doi.org/10.1080/17439884.2020.1798995.
[23] Frey CB, Osborne MA. The future of employment: how susceptible are jobs to computerisation? Technol Forecast Soc Change 2017;114:254–80. https://doi.org/10.1016/j.techfore.2016.08.019.
[24] Autor DH, Levy F, Murnane RJ. The skill content of recent technological change: an empirical exploration. Q J Econ 2003;118(4):1279–333. https://doi.org/10.1162/003355303322552801.
[25] Goos M, Manning A, Salomons A. Explaining job polarization: routine-biased technological change and offshoring. Am Econ Rev 2014;104(8):2509–26. https://doi.org/10.1257/aer.104.8.2509.
[26] Brynjolfsson E, Mitchell T. What can machine learning do? Workforce implications. Science 2017;358(6370):1530–4. https://doi.org/10.1126/science.aap8062.
[27] Raisch S, Krakowski S. Artificial intelligence and management: the automation–augmentation paradox. Acad Manag Rev 2021;46(1):192–210. https://doi.org/10.5465/amr.2018.0072.
[28] Katz LF, Murphy KM. Changes in relative wages, 1963–1987: supply and demand factors. Q J Econ 1992;107(1):35–78. https://doi.org/10.2307/2118323.
[29] Autor D. The labor market impacts of technological change: from unbridled enthusiasm to qualified optimism to vast uncertainty (NBER working paper no. w30074). National Bureau of Economic Research; 2022. https://doi.org/10.3386/w30074.
[30] Crompton H, Burke D. Artificial intelligence in higher education: the state of the field. Int J Educ Technol High Educ 2023;20(1):22. https://doi.org/10.1186/s41239-023-00392-8.
[31] Lave J, Wenger E. Situated learning: legitimate peripheral participation. Cambridge University Press; 1991. https://doi.org/10.1017/CBO9780511815355.
[32] Vygotsky LS. Mind in society: the development of higher psychological processes. Harvard University Press; 1978.
[33] Bao H, Sun M, Teplitskiy M. Where there's a will there's a way: chatGPT is used more for science in countries where it is prohibited. Quant Sci Stud 2025;6:716–31. https://doi.org/10.1162/qss_a_00368.
[34] Yu H. Reflection on whether Chat GPT should be banned by academia from the perspective of education and teaching. Front Psychol 2023;14. https://doi.org/10.3389/fpsyg.2023.1181712. Article 1181712.
[35] Gillespie N, Lockey S, Ward T, Macdade A, Hassed G. Trust, attitudes and use of artificial intelligence: a global study 2025. The University of Melbourne & Kpmg; 2025. https://assets.kpmg.com/content/dam/kpmgsites/xx/pdf/2025/05/trust-attitudes-and-use-of-ai-global-report.pdf.
[36] Corbett C, Hill C. Solving the equation: the variables for women's success in engineering and computing. American Association of University Women; 2015.
[37] Benjamin R. Race after technology: abolitionist tools for the new JIM code. Polity Press; 2019.
[38] Chan CKY. A comprehensive AI policy education framework for university teaching and learning. Int J Educ Technol High Educ 2023;20(1):38. https://doi.org/10.1186/s41239-023-00408-3. Article.
[39] Young E, Wajcman J, Sprejer L. Mind the gender gap: inequalities in the emergent professions of artificial intelligence (AI) and data science. New Technol Work Employ 2023;38(3):391–414. https://doi.org/10.1111/ntwe.12278.
[40] O'neil C. Weapons of math destruction: how big data increases inequality and threatens democracy. Crown; 2017.
[41] Bao L, Huang D, Lin C. Can artificial intelligence improve gender equality? Evidence from a natural experiment. Manag Sci 2024. https://doi.org/10.1287/mnsc.2022.02787. advance online publication.
[42] Creswell JW, Plano Clark VL. Revisiting mixed methods research designs twenty years later. In: Hesse-Biber SN, Johnson RB, editors. The sage handbook of mixed methods research design. Sage; 2023. p. 21–36. https://doi.org/10.4135/9781529614572.n6.
[43] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15(2):155–63. https://doi.org/10.1016/j.jcm.2016.02.012.
[44] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86(2):420–8. https://doi.org/10.1037/0033-2909.86.2.420.
[45] Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6(4):284. https://psycnet.apa.org/doi/10.1037/1040-3590.6.4.284.
[46] Ha E, Ognyanova K, Singh V, Uslu A. National AI opinion monitor (NAIOM). Rutgers University; 2025. https://naiom.net/public-reports/NAIOM%20Report%2003%20AI%20Use%20and%20Attitudes%20in%20NJ.pdf.
[47] Luo T. How does GenAI affect trust in teacher-student relationships? Insights from students' assessment experiences. Teach High Educ 2024. https://doi.org/10.1080/13562517.2024.2304892. Advance online publication.
[48] Watson CE, Rainie L. The AI challenge: how college faculty assess the present and future of higher education. AAC&U & Elon University Imagining the Digital Future Center; 2026. https://imaginingthedigitalfuture.org/wp-content/uploads/2026/01/Elon-AACU-faculty-AI-survey-full-report-1-21-26.pdf.
[49] Kestin G, Miller K, Klales A, Milbourne T, Ponti G. AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting. Sci Rep 2025;15(1):17458. https://doi.org/10.1038/s41598-025-97652-6.
[50] Gu J, Yan Z. Effects of GenAI interventions on student academic performance: a meta-analysis. Comput Educ: Artif Intell 2025;8:100400. https://doi.org/10.1016/j.caeai.2025.100400. Article.
[51] Cengage Group. (2024). *2024 graduate employability report*. https://www.cengagegroup.com/news/press-releases/2024/cengage-group-2024-employability-report/.
[52] Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Routledge; 1988. https://doi.org/10.4324/9780203771587.
[53] National Center for Education Statistics. Table 318.30. bachelor's, master's, and doctor's degrees conferred by postsecondary institutions, by sex of student and field of study: academic year 2021–22. U.S. Department of Education; 2023. https://nces.ed.gov/programs/digest/d22/tables/dt22_318.30.asp.
[54] National Center for Education Statistics. Table 505.06. number and percentage distribution of 25- to 64-year-old bachelor's degree holders, percentage of degree holders among all 25- to 64-year-olds, and unemployment rates and median annual earnings of 25- to 64-year-old bachelor's degree holders, by age group, field of study, and science, technology, engineering, or mathematics (STEM) status of field: 2022. U.S. Department of Education; 2023. https://nces.ed.gov/programs/digest/d22/tables/dt22_505.06.asp.