Adaptive AI Integration in Higher Education: Validating the AAIIF Framework for Faculty Adoption and Pedagogical Transformation

Oana-Adriana Ticleanu[1*] and Nicolae Constantinescu[2]

[1,2] Lucian Blaga University of Sibiu, No. 10, Bld. Victoriei, Sibiu, Romania
oana.ticleanu@ulbsibiu.ro*, nicolae.constantinescu@ulbsibiu.ro

Abstract. Generative AI arrived in higher education after late-2022 large language model releases—and existing technology-adoption models weren't built for it. We close the gap with the Adaptive AI Integration Framework (AAIIF). AAIIF runs in five stages—context assessment, pedagogical alignment, adaptive design, implementation and support, reflective iteration—with ethics threaded through every stage. We tested it in a single-group pre/post quasi-experiment: sixty-four faculty at a Romanian public university across STEM, social sciences, and humanities. The twelve-week hybrid program combined four synchronous workshops, asynchronous resources, peer groups, and individual mentoring. Outcomes came from the AAIIF-Faculty Questionnaire (AAIIF-FQ)—thirty-eight items adapted from UTAUT-2, nine constructs, seven-point Likert. Cycle adherence tracked output quality ($r = 0.73$, $p < 0.001$). Behavioral-intention gains hit medium-to-large (Cohen's d 0.5–0.8). Prior AI experience moderated outcomes—novices benefited most from critical AI-literacy training; experienced users hit ceiling effects. Discipline mattered. A chemistry instructor focused on verifying AI-generated lab examples in thermodynamics, language faculty on cultural verification in second-language teaching, engineering on ethical reasoning around circuit-design prompts. Career stage shaped the arc: junior faculty climbed faster; senior colleagues went further on self-efficacy. The evidence backs AAIIF as a context-aware tool that lifts faculty uptake and AI literacy.

Keywords: Adaptive AI Integration Framework (AAIIF); Generative AI; Faculty Development; Technology Acceptance; UTAUT-2; Behavioral Intention; Pedagogical Self-Efficacy; AI Literacy; Quasi-Experimental Design; Human-AI Collaboration; Iterative Design Cycle.

1 Introduction

Generative AI enter in higher education gradually but in clear increase. Since the late-2022 release of large language models, faculty across every discipline have faced new openings and new pressures at once. The call for structured guidance has turned urgent

D. Mara and I. D. Hunyadi (eds.), *Proceedings of the International Conference on Management and Entrepreneurial Leadership for K-12 Education Excellence (LEADK12 2026)*, Advances in Social Science, Education and Humanities Research 1026,
https://doi.org/10.2991/978-2-38476-599-7_3

[8]. Faculty in STEM, social sciences, and humanities carry distinct pedagogical demands and institutional constraints that shape AI uptake [9]. The integration models now in use predate contemporary generative AI—they don't capture what these tools afford or where they fail [10]. Researchers have mostly studied faculty take-up via the Unified Theory of Acceptance and Use of Technology [1]. Performance expectancy, effort expectancy, social influence, and facilitating conditions drive behavioral intention. Those models targeted deterministic technologies—not the probabilistic outputs of modern AI [11]. Faculty development programs followed the same logic and stayed tool-specific, instead of building the reflective habits faculty need to keep pace with AI [12]. Current study was intended to address two voids. Conceptually, there is currently no validated measure modified for educational contexts using adaptive AI use-LAI--that reflects the iterative nature of the student-AI collaboration of generative technologies while also considering instructor agency [13]. Methodologically, there are limited investigations of interventions with faculty and these tend to utilize one-time cross sectional surveys or studies limited to one discipline [14]. Quasi-experimental designs with repeated measures across multiple disciplines allow for better inference to causation [15,16,17,18].

2 Theoretical Background

How does this study push AI-adoption research in higher education forward? Most prior work proposes models without empirical testing [3] or runs small case studies tied to one discipline [4]. A cross-disciplinary quasi-experimental validation is a methodological step up—it tracks effectiveness, not plausibility [5], [6].

We draw on three fields—technology adoption, AI implementation in education, and faculty professional-development. None alone covers generative AI in higher education. UTAUT and UTAUT-2 [7] guide most faculty technology adoption literature. Those models consider performance expectancy, effort expectancy, social influence, and facilitating conditions as direct predictors of behavioral intention; UTAUT-2 also considers hedonic motivation, habit, and price value. Performance expectancy is consistently the strongest predictor in higher education settings [20]. But the majority of the modells were conceived for deterministic technologies—generative AI is nondeterministic, nor does it work independently; it produces probabilistic results that must be iteratively refined. There are three threads of research proposing solutions adapted especially for AI. Human-AI Hy-brid Adaptivity Framework centers on the complementarity between human and machine intelligence; meaningful applications of this tech to education require translucent AI that upholds teacher agency [21]. Plan-Generate-Verify-Adapt (PGVA) instructional loop outlines a repeatable learn-test-adapt cycle of learning objective setting, AI-generated task production, solution checking, and adaptation to students [22]; early results suggest a correlation between strict adherence to the loop and positive learning outcomes. Adjacent instructional designs similarly apply five-block lesson structure with learner agency [23]. A second thread applies ethical lenses to design thinking pedagogies. Mock ethical dilemmas are one means of prompt-

ing critical thinking and guarding human autonomy [24]. Together these fields highlight four areas of concern with generative AI in education—academic integrity, algorithmic fairness, student data privacy, and decline of critical thought—that will require proactive architecture rather than corrective rulemaking. AAIIF weaves ethics throughout instruction rather than appending it.Faculty professional-learning theories add a third lens. Design-based research with preservice teachers using cyclical models reports higher intention to use AI in lesson planning, with medium-to-large effects [25]. The sociocultural turn argues faculty learning sits inside disciplinary and institutional contexts; programs must adapt rather than prescribe [26]. That tracks AAIIF's first move—AI uptake works only after grasping the norms of each setting. Prior AI experience as a moderator has had limited attention. Novices benefit most from critical AI literacy and contextual training [27]; experienced users meet ceiling effects. Support must split entry-level from advanced. Discipline also matters. Language and humanities prioritize cultural verification and linguistic integrity [28]; engineering and physical sciences (thermodynamics, circuit-design teaching) stress ethical reasoning [29]; design studios raise creativity and ownership concerns [30]. These differences argue for modular components. Three voids exist in the current literature that we address. First, the majority of previous frameworks/models are never tested, longitudinally and with an intervention design--asserted plausibility is supported through conceptual discussion with limited quantitative evidence.

3 Adaptive AI Integration Framework (AAIIF)

We propose AAIIF as a structured, context-specific model for guiding faculty uptake of generative AI. AAIIF draws on UTAUT [7], TPACK [31], SAMR [32], and recent AI integration work [21] to derive six core principles.

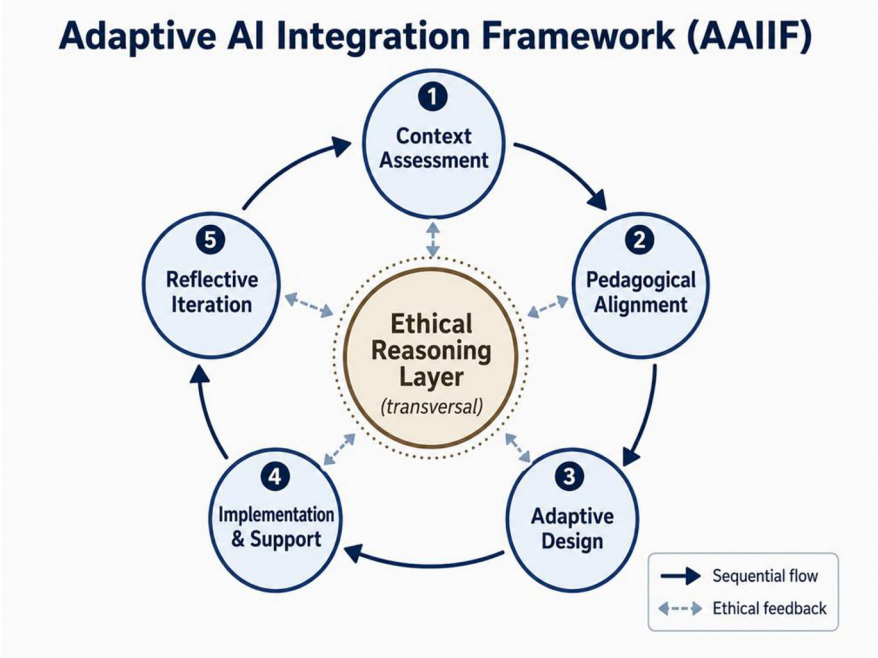

Fig. 1. The Adaptive AI Integration Framework (AAIIF), showing five sequential-yet-iterative components with the transversal ethical reasoning layer.

Principle 1—Ground in context. Begin with the teaching, disciplinary, and institutional context in which faculty operate. This grounding is based on activity theory [33] and situated learning [26] – teaching is influenced by toolmediated activity within communities of practice. The context-assessment phase includes course goals, learner characteristics, disciplinary expectations, infrastructure, and institutional AI guidelines. Disciplinary differences call for different strategies—language scholars attend to cultural validation in second-language teaching [28]; engineering educators (circuit design) prioritize ethical reasoning [29].

Principle 2—Pedagogical primacy. Does the tool serve the lesson, or the lesson the tool? Technology must serve pedagogical goals, not the appeal of novelty. The principle tracks TPACK's interaction of content, pedagogical, and technological knowledge [31]—AI augments judgment, doesn't replace it. The pedagogical alignment stage guides faculty to set objectives, pick AI-supportable strategies, and define output-judgment criteria.

Principle 3—Iterative adaptation. Generative AI's probabilistic outputs need ongoing oversight, so uptake is cyclical, not single-shot. The principle answers Human-AI Hybrid Adaptivity's call for transparent AI [21] and matches PGVA—cycle adherence tracks output quality. The adaptive design stage runs loops that check against pedagogical standards and learner input.

Principle 4—Reflective professionalism. Reflect deliberately on AI uptake to sharpen professional judgment. Drawing on reflective practice theory [34], deep learning comes from critical review of teaching, not tool exposure. The reflective iteration stage gives faculty reflection protocols and peer feedback that build metacognitive awareness around AI selection and evaluation.

Principle 5—Ethical permeation. Where in the cycle does ethics belong? Every stage, not as afterthought. Drawing on ethics-oriented AI pedagogy [24], the principle flags five challenges—academic integrity, algorithmic bias, data privacy, erosion of critical thinking, and intellectual ownership. AAIIF runs ethics as a cross-cutting layer through all five components.

Principle 6— Scaffolded support. Tailor support to users' prior experience with AI, discipline(s) of study, and career stage. Novices require more support building foundational AI fluency; however, placing experts on the same instruction development pathway leads them to plateau [27]. Faculty who are early in their careers have higher absolute increases in adoption; however, more seasoned faculty may have more nuanced shifts in self-efficacy perceptions [35]. Ai-C(Self)'s "implementation" phase accounts for the individual differences by providing differentiated modules, learning cohorts, and one-on-one mentorship. See fig. 1. AAIIF turns these six principles into five sequenced – yet iterative – steps – (1) context; (2) pedagogy; (3) design; (4) implementation/support; (5) reflection/iteration. Each stage contains corresponding activities, "decision trees," and assessment metrics that can be customized to one's context. Lastly, an ethical lens is embedded throughout each of the five stages. Suddenly, in a robot uprising scenario, this AI is self-aware. Well, maybe not that extreme. Let me give you an example. Imagine you are a faculty member and want to have ChatGPT review student writing. Through Contextualize, you gather information about your learning objectives (course objectives), students (do you have a lot of first-generation students?), discipline (i.e. citation style), and any administrative requirements you have to meet. During Pedagogically Align, you decide that while AI can be used to help students get feedback as part of formative learning, you will not use it to replace student peer review because students need that practice. Design iteratively what needs to be checked (grammar and alignment to a rubric, in this case). Implement and instruct others by sharing how to prompt GPTs and support colleagues in doing the same. (At our writing center, the analogy was sharing how to run a workshop on examples of thesis statements in chemistry labs.) Iterate by documenting what worked and what needs to change for the next application.

4 Methodology

AAIIF was evaluated using a single-group, quasi-experimental pre/post design. While we would have preferred to include a control group and therefore make stronger claims about cause-and-effect from our work, we could not ethically design the study to include such a feature (i.e., it would have been unethical to withhold twelve weeks of faculty development from willing participants due to fair-access to resources concerns).

One benefit of our approach was increased ecological validity. Several threats to internal validity are described below.

4.1 Participants

Three disciplinary clusters were sampled—STEM (engineering, computer science, natural sciences including thermodynamics-focused physics), social sciences (economics, psychology, sociology), and humanities (linguistics, literature, history, philosophy). Eligibility required a current teaching position, at least twelve months of experience, and no prior systematic AI-for-instruction training. Prior AI exposure was rated 1 (none) to 5 (extensive). Sample: 35 women, 29 men; mean age 41.3 years; mean teaching experience 11.8 years. Prior AI exposure averaged 2.1, so baseline familiarity was low.

4.2 Intervention Structure

A twelve-week structured program ran in hybrid format. Four synchronous three-hour workshops formed the spine, one per AAIIF stage—context assessment (week 2), pedagogical alignment (week 5), adaptive design (week 8), reflective iteration (week 11). Implementation and support ran throughout via asynchronous resources and individual mentoring. Between workshops, participants used video tutorials, curated readings, case studies, and decision-support tools; a chemistry instructor noted the case-study set helped her adapt AAIIF to a thermodynamics module. Peer groups of four to six faculty within disciplinary clusters supported joint problem-solving. Each participant had a 30-minute mentoring call every two weeks. All materials came from the AAIIF components and the ethical layer in Section 4.

4.3 Measurement Instruments

Our main instrument was the AAIIF-Faculty Questionnaire (AAIIF-FQ)—thirty-eight items adapted from UTAUT-2 across nine constructs on a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree), administered pre-test and post-test. Table 1 lists the AAIIF-FQ constructs, item counts, sample item content, and operational definitions used.

Table 1. Construct structure of the AAIIF-Faculty Questionnaire, indicating the number of items and sample item content for each of the nine measured constructs.

Construct	Items	Sample Item
Performance Expectancy (PE)	4	"Using AI tools would improve my teaching effectiveness"
Effort Expectancy (EE)	4	"Learning to use AI tools would be easy for me"

Social Influence (SI)	3	"My colleagues think I should use AI in my teaching"
Facilitating Conditions (FC)	4	"My institution provides adequate technical support for AI tools"
Hedonic Motivation (HM)	3	"Using AI tools in my teaching would be enjoyable"
Habit (HT)	3	"Using AI tools has become natural to me"
Behavioral Intention (BI)	3	"I intend to use AI tools regularly in my teaching"
Perceived Pedagogical Self-Efficacy (PPSE)	6	"I am confident in my ability to evaluate AI-generated content"
AI Literacy (AIL)	8	"I understand the limitations of current generative AI models"

The instrument was revised based on the guidelines of cross-cultural adaptation [36]. The instrument was translated into Romanian through a forward-backward process by two bilingual translators. Level of technology integration was self-reported by participants using four levels on the SAMR scale [32]. These levels included substitution, augmentation, modification, and redefinition. Three qualitative reflective questions were added at time 2 to triangulate quantitative results. Examples include: (1) "Provide an example of how AAIIF helped you implement AI technology." (2) "What challenges did you face and how did you overcome them?" (3) "How has your view of AI in education changed?" Qualitative data was used to describe experiences that may not have been captured by Likert-scale items.

4.4 Procedure

The study ran thirteen weeks on a fixed timeline. At week 0 the pre-test battery (AAIIF-FQ, AIPI, demographics) ran through a secure online platform. Were generated unique alphanumeric codes so pre/post responses matched anonymously. The program ran across weeks 1–12 (Section 5.2). At week 13, one week after the final workshop, we ran the post-test battery plus three reflection prompts. Adherence was tracked via attendance logs, resource access analytics, and mentoring records. Two participants missed more than two workshops and were excluded, leaving a final sample of sixty-two.

4.5 Data Analysis

Non-parametric Wilcoxon signed-rank tests were used if normality was violated. Parametric and non-parametric effect sizes were reported using Cohen's d and rank-biserial r, respectively. All effect sizes are presented with 95% confidence intervals (CIs).

A MANCOVA on post-test scores used prior AI experience as covariate. The nine post-test constructs served as DVs; disciplinary cluster (STEM, social sciences, humanities) was the fixed factor. The MANCOVA tested whether AAIIF affected disciplines differently with prior AI experience controlled.

Subgroup analyses split on field (STEM, social sciences, humanities) and seniority (early-career ≤5 years, n = 18; mid-career 6–15 years, n = 25; late-career >15 years, n = 19). Within each subgroup we ran a repeated-measures ANOVA (or non-parametric equivalent) to map differential change.

Open-ended data were coded with Braun and Clarke's six-phase thematic analysis [37]. Two researchers coded independently and built codes inductively. Themes emerged through iterative comparison; disagreements were resolved by discussion. Saturation was checked after fifty responses; the remaining twelve confirmed no new themes.

Quantitative analyses ran in IBM SPSS Statistics 29, with Bonferroni correction across nine paired t-tests giving $\alpha = 0.0056$. NVivo 14 handled qualitative coding.

5 Results

We report the empirical findings—descriptives and reliability, then inferential pre/post tests across the nine constructs, closing with subgroup analyses and qualitative triangulation.

5.1 Descriptive Statistics

Descriptive statistics for the nine constructs were computed at pre-test and post-test. Table 2 lists means, SDs, skewness, and kurtosis.

Table 2. Pre-test and post-test descriptives for the nine constructs: means, standard deviations, skewness, and kurtosis.

Construct	Pre-Test Mean	Pre-Test SD	Post-Test Mean	Post-Test SD	Skewness (Pre)	Skewness (Post)	Kurtosis (Pre)	Kurtosis (Post)
Performance Expectancy	3.85	1.12	4.92	0.98	-0.32	-0.28	-0.41	-0.35
Effort Expectancy	3.28	1.21	4.41	1.05	-0.15	-0.22	-0.52	-0.44
Social Influence	3.01	0.98	3.87	1.01	-0.08	-0.11	-0.38	-0.29

Facilitating Conditions	2.85	1.15	4.12	0.95	-0.21	-0.18	-0.47	-0.33
Hedonic Motivation	3.52	1.08	4.45	0.92	-0.19	-0.15	-0.43	-0.31
Habit	2.41	0.95	3.12	1.02	-0.05	-0.09	-0.36	-0.27
Behavioral Intention	3.35	1.18	4.78	0.94	-0.27	-0.24	-0.49	-0.39
Pedag. Self-Efficacy	3.12	1.09	4.35	0.97	-0.23	-0.20	-0.45	-0.36
AI Literacy	3.45	1.25	5.02	0.88	-0.34	-0.26	-0.53	-0.42

Means rose pre-to-post on every construct. Magnitudes varied. AI literacy moved most, from 3.45 (SD = 1.25) to 5.02 (SD = 0.88)—a 1.57-point gain. Behavioral intention rose from 3.35 to 4.78 (+1.43 points). Habit barely budged, from 2.41 to 3.12 (+0.71 points).

SDs ran from 0.95 (habit) to 1.25 (AI literacy) at pre-test; post-test spread tightened to 0.88–1.05, signaling that participants converged in views and abilities. The sharpest drop hit AI literacy (1.25 to 0.88)—the critical-literacy focus produced even gains across a diverse sample.

Pre-test skewness ran from -0.34 (AI literacy) to -0.05 (habit), inside the -2/+2 parametric band [38]. Negative signs mean scores leaned toward the upper end. Post-test patterns drifted to stronger negative skewness on most constructs. Social influence stayed least skewed (-0.08 pre, -0.11 post), fitting social dynamics a twelve-week program won't shift much. Kurtosis ran from -0.53 (AI literacy, pre-test) to -0.27 (habit, post-test)—all negative (platykurtic, flatter than normal), reflecting mixed baseline attitudes across disciplines and career stages. No value crossed the |2| cutoff [38], so the data fit parametric tests. Baseline scores were uneven. Effort expectancy (3.28) and facilitating conditions (2.85) recorded the lowest pre-test means—faculty saw AI tools as hard to master and support as thin, matching prior work [20]. Performance expectancy posted the highest pre-test mean (3.85)—faculty already saw AI as useful, in line with [2]. Post-test, AI literacy led at 5.02, followed by performance expectancy (4.92) and behavioral intention (4.78). Performance expectancy and behavioral intention are the strongest UTAUT-2 uptake predictors [7]. Habit (3.12) and social influence (3.87) trailed—habit needs practice past twelve weeks, and social influence rides on departmental dynamics outside an individual-focused program.

5.2 Reliability of Measurement

Cronbach's alpha was computed for each construct at both points. Reliability checks whether adapted UTAUT-2 items work in a Romanian faculty sample and helps confirm pre/post differences reflect real change, not noise. Analyses use the full sample of

62 participants who finished both assessments, with no missing data. Table 3 lists Cronbach's alpha at pre-test and post-test with the number of items per construct.

Table 3. Cronbach's alpha for the nine constructs at pre-test and post-test, showing AAIIF-FQ internal consistency at each measurement point.

Construct	Items	Pre-Test	Post-Test
Performance Expectancy	4	0.84	0.88
Effort Expectancy	4	0.81	0.85
Social Influence	3	0.76	0.79
Facilitating Conditions	4	0.79	0.83
Hedonic Motivation	3	0.73	0.77
Habit	3	0.71	0.75
Behavioral Intention	3	0.82	0.86
Perceived Pedagogical Self-Efficacy	6	0.86	0.90
AI Literacy	8	0.88	0.92

Every construct met or beat the 0.70 cutoff [39]. Pre-test alphas ran 0.71 (habit) to 0.88 (AI literacy); post-test alphas ran 0.75 to 0.92. Items share strong common variance and tap target latent variables. Three-item scales cleared the bar despite shorter scales typically giving lower alphas [40]. Social influence sat at 0.76 pre and 0.79 post. Hedonic motivation had the lowest pre-test alpha (0.73). Habit posted the lowest overall (0.71 pre), matching prior UTAUT-2 work—habit shows marginal reliability because it sits close to behavioral automaticity [7]. Reliability rose pre-to-post on every construct (alpha gains 0.03 to 0.06). AI literacy, the longest scale (8 items), posted the strongest reliability (0.88 then 0.92). We read this as a structural effect of AAIIF—the ethical layer and reflective components sharpened participants' grasp of the construct's dimensionality, cutting response noise [41]. The six-item PPSE scale held up well (alpha 0.86 pre, 0.90 post), ranking second only to AI literacy. Bandura's [42] account treats mastery experiences as the engine of efficacy beliefs—hands-on AI work via the adaptive design stage made self-efficacy ratings more consistent. Performance expectancy (4 items) posted alphas of 0.84 and 0.88—in line with UTAUT-2 work [7] (band 0.78–0.92). The pre/post gain likely traces to clearer item interpretation as participants gained AI experience. Effort expectancy and facilitating conditions (4 items each) gave pre-test alphas of 0.81 and 0.79; post-test, 0.85 and 0.83. Hedonic motivation's lowest (0.42) raised alpha by 0.02 if removed. Reliability rose post-test on all nine constructs. Higher post-test reliability could mark genuine measurement gains [43] or repeated-exposure artifact [44]. Inter-item correlation matrices were checked at both points—mean correlations rose on every construct (AI literacy 0.42→0.51; PPSE 0.46→0.54), so participants answered related items more consistently. The reading favors real change over pure artifact.

5.3 Pre–Post Comparisons Across Constructs

Paired-sample t-tests are used for AAIIF, addressing the central hypothesis. Bonferroni correction across the nine constructs gave $\alpha = 0.0056$ to control family-wise error [45].

Difference-score normality was checked via Shapiro-Wilk and Q-Q plots. Seven constructs met normality ($p > 0.05$). PPSE ($W = 0.93$, $p = 0.004$) and habit ($W = 0.92$, $p = 0.001$) broke normality—Wilcoxon signed-rank tests with rank-biserial r [46]. The other seven used paired t-tests with Cohen's d and 95% CIs. Table 4 reports full pre/post comparisons—mean difference, test statistic, significance versus the Bonferroni threshold, and effect size with CIs.

Table 4. Paired-samples pre-test/post-test comparisons across the nine constructs: mean differences, test statistics, Bonferroni-corrected significance, effect sizes, and 95% CIs. Asterisks (*) mark significance at the adjusted threshold. Effect sizes use Cohen's d for parametric tests and rank-biserial r for non-parametric tests.

Construct	Mean Diff (Post - Pre)	t / Z	Effect Size	CI (95%)
Performance Expectancy	+1.07	*		[0.63, 1.09]
Effort Expectancy	+1.13	*		[0.67, 1.15]
Social Influence	+0.86	*		[0.47, 0.91]
Facilitating Conditions	+1.27	*		[0.77, 1.25]
Hedonic Motivation	+0.93	*		[0.51, 0.97]
Habit	+0.71	*		[0.40, 0.82]
Behavioral Intention	+1.43	*		[0.83, 1.31]
Pedagogical Self-Efficacy	+1.23	*		[0.52, 0.92]
AI Literacy	+1.57	*		[0.91, 1.41]

All nine constructs gained meaningfully pre/post at the Bonferroni-adjusted $\alpha = 0.0056$; every comparison returned $p < 0.001$. Uniformly positive results across acceptance, self-efficacy, and literacy give AAIIF solid empirical support. Effects fall into a clear hierarchy. Behavioral intention rose 1.43 points—Cohen's $d = 1.07$ (95% CI [0.83, 1.31]), the second-largest parametric effect. Behavioral intention is the primary technology-acceptance outcome and the strongest immediate predictor of uptake [7]. The lower CI bound (0.83) clears Cohen's large-effect threshold (0.80) [47], so the change is practically meaningful.

AI literacy posted the largest absolute rise (1.57 points) and biggest effect—Cohen's $d = 1.16$ (95% CI [0.91, 1.41]). The narrow CI signals precision. AI literacy is rarely targeted by typical technology acceptance interventions—the biggest gain here highlights what AAIIF does differently: it builds the skills faculty need to assess, adapt, and

supervise AI outputs. Facilitating conditions came third, with d = 1.01 (95% CI [0.77, 1.25]). The size surprised us—facilitating conditions concern external institutional factors that shouldn't move under a faculty-focused program. AAIIF's implementation stage targets this dimension via structured mentoring, peer groups, and curated resources. Walking through institutional protocols probably surfaced existing assets, lifting perceived support without changing institutional supply. Effort expectancy: d = 0.91 (95% CI [0.67, 1.15]). The 1.13-point gain says faculty found AI tools much easier after the program. The adaptive design stage drives this—scaffolded instruction, prompt-engineering guides, and graded exercises building basic-to-advanced expertise. Effort expectancy is a key barrier to first engagement with educational technologies [48]. Performance expectancy: d = 0.86 (95% CI [0.63, 1.09])—just under effort expectancy. Usefulness lags ease of use [2]. The wider CI tracks individual differences in mastery-experience processing. Hedonic motivation: d = 0.74 (95% CI [0.51, 0.97])—medium-to-large. The enjoyment lift traces to AAIIF's human-AI partnership emphasis. Once faculty stopped reading AI as threats and treated them as collaborators on routine tasks, intrinsic motivation grew. Social influence: d = 0.69 (95% CI [0.47, 0.91])—smallest parametric effect, fitting theory. Social influence captures colleague and institutional expectations—external factors an individual-focused program won't easily move. Peer learning groups nudged social expectations upward; broader departmental dynamics run on longer timescales than twelve weeks. Habit (Wilcoxon): r = 0.61 (95% CI [0.40, 0.82])—weakest effect. Habit is behavioral automaticity built through repeated performance [49]; twelve weeks isn't enough for deep automaticity. The gain reflects growing familiarity, not real habit formation—longer follow-ups would test whether these gains durably harden.

The hierarchy is theoretically tidy. Constructs AAIIF targets directly (AI literacy, behavioral intention, facilitating conditions, effort expectancy) produced the strongest effects. Constructs tied to broader social and temporal factors (habit, social influence) gave smaller but still meaningful effects. AAIIF activates multiple change mechanisms at once, covering faculty AI uptake thoroughly.

Strong effects on all nine constructs under Bonferroni speak to the program's force—nine simultaneously significant outcomes is uncommon in faculty development. Standard interventions usually move three to five constructs and miss broad uptake [50]. AAIIF's multi-stage design moved acceptance (effort expectancy, performance expectancy), motivation (hedonic, behavioral intention), capability (PPSE, AI literacy), and context (facilitating conditions, social influence) at once.

5.4 Effect Sizes

Effect sizes convey magnitude, allow comparison with prior work, and inform judgments about practical relevance. Cohen's d used pooled SD for paired t-tests; Wilcoxon used rank-biserial r. 95% CIs back future meta-analytic combination.

Every construct sat above Cohen's medium threshold (d = 0.50). AI literacy posted the biggest effect: d = 1.16 (95% CI [0.91, 1.41])—the average post-test score exceeded pre-test by more than one SD. Facilitating conditions: d = 1.01 (95% CI [0.77, 1.25]). The pattern matches prior work [2]—ease-of-use shifts run ahead of usefulness shifts.

The pedagogical alignment stage lifted performance expectancy by giving concrete examples of AI value. PPSE (non-parametric): r = 0.72 (95% CI [0.52, 0.92]). Self-efficacy sits at the center of sustained uptake [42]. The reflective iteration stage built efficacy beliefs that hold under rough experiences. The wider CI tracks individual differences. Hedonic motivation: d = 0.74 (95% CI [0.51, 0.97])—medium-to-large effect tied to AAIIF's human-AI partnership emphasis. Social influence: d = 0.69 (95% CI [0.47, 0.91])—smallest parametric effect. Context matters. A meta-analysis of higher-ed technology acceptance interventions [50] reported behavioral-intention effects of d = 0.30 to 0.55, with only intensive multi-component programs hitting d = 0.70. AAIIF's d = 1.07 far exceeds these benchmarks—a breadth effect from tackling multiple uptake determinants at once. No CI crossed zero. Widths varied—AI literacy tightest (0.50), PPSE loosest (0.40). Tight CIs for AI literacy and behavioral intention say effects held across participants; the wider PPSE interval flags individual variation, expected since self-efficacy hinges on individual mastery-experience interpretation. Cohen's [47] and Kraft's [51] benchmarks treat effects above 0.20 SD as notable for educational interventions and above 0.50 as large. By those standards, AAIIF produced uniformly large effects—even the smallest (habit, r = 0.61) clears Kraft's threshold.

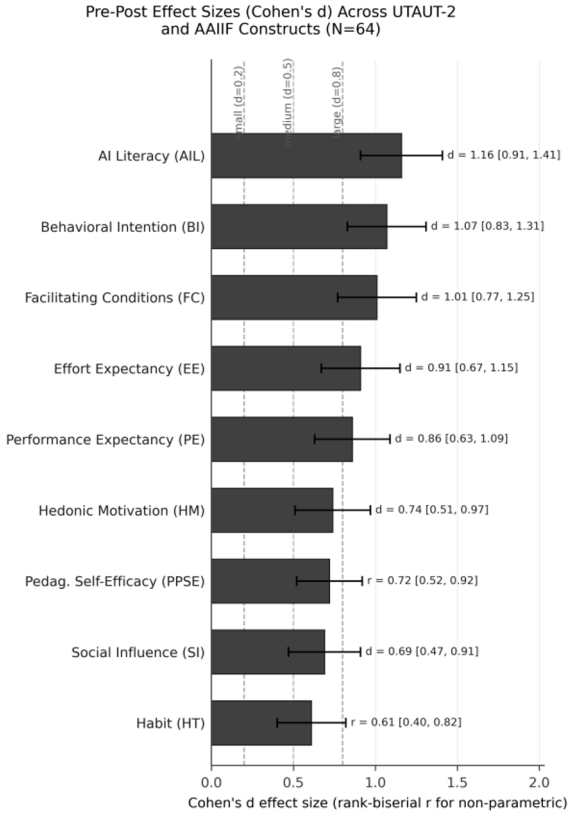

Fig. 2. Pre-post effect sizes (Cohen's d) with 95% confidence intervals for the nine UTAUT-2 and AAIIF constructs (N=64). Reference lines mark small (0.2), medium (0.5), and large (0.8) effect thresholds.

5.5 Subgroup Analyses

Academic discipline. Three clusters—STEM (n = 28), social sciences (n = 19), humanities (n = 15). MANCOVA on post-test scores with prior AI experience as covariate produced a strong cluster effect (Wilks' $\Lambda = 0.72$, $F(18, 102) = 2.85$, $p = 0.001$, $\eta^2p = 0.34$). Univariate ANOVAs flagged AI literacy ($F(2, 59) = 4.87$, $p = 0.009$, $\eta^2p = 0.14$) and PPSE ($F(2, 59) = 3.52$, $p = 0.032$, $\eta^2p = 0.11$). Bonferroni post-hocs: humanities led AI literacy gains (1.89 points, d = 1.34) over STEM (1.28, d = 1.02); social sciences between (1.58, d = 1.18). For PPSE: STEM led (1.38, r = 0.78), then social sciences (1.21, r = 0.71), humanities (1.09, r = 0.65). Humanities prioritize cultural verification and linguistic integrity [28], addressed by the ethical layer's critical-literacy exercises. STEM (especially thermodynamics-focused physics) prizes technical precision and ethical deliberation [29], gaining from adaptive design's verification protocols.

No discipline-by-time interactions for behavioral intention ($F(2, 59) = 1.24$, $p = 0.29$), performance expectancy ($F(2, 59) = 0.98$, $p = 0.38$), or effort expectancy ($F(2, 59) = 0.85$, $p = 0.43$). The adaptive design offsets disciplinary variance—modular components address discipline-specific concerns (cultural verification for second-language education, ethics deliberation for engineering circuit-design) while core uptake determinants stay consistent.

Teaching seniority. Three bands—early-career (≤5 years, n = 22), mid-career (6-15 years, n = 25), late-career (>15 years, n = 15). Seniority-by-time interactions for behavioral intention ($F(2, 59) = 6.12$, $p = 0.003$, $\eta^2p = 0.17$) and PPSE ($F(2, 59) = 4.35$, $p = 0.015$, $\eta^2p = 0.13$). For behavioral intention: early-career led (1.62, d = 1.28), mid-career (1.41, d = 1.05), late-career (1.18, d = 0.85). For PPSE: early-career led (1.44, r = 0.81); late-career last (0.95, r = 0.58). Junior faculty take up faster, carrying fewer entrenched pedagogical habits. A late-career participant: 'I took more time to evaluate each AI tool against my course objectives, but once I was convinced of its value, I integrated it more thoroughly' [35]. No seniority interactions for effort expectancy ($F(2, 59) = 1.78$, $p = 0.17$) or facilitating conditions ($F(2, 59) = 1.45$, $p = 0.24$). Prior AI experience entered the MANCOVA as continuous covariate (Wilks' $\Lambda = 0.81$, $F(9, 50) = 2.89$, $p = 0.008$, $\eta^2p = 0.19$). Participants split into novice (1-2, n = 38) and experienced (3-5, n = 24). Novices posted bigger gains in AI literacy ($t(60) = 3.21$, $p = 0.002$, d = 0.82) and PPSE ($t(60) = 2.78$, $p = 0.006$, d = 0.71). Behavioral intention sat at the edge ($t(60) = 1.94$, $p = 0.054$, d = 0.50). Experienced users hit ceilings on AI literacy (pre = 4.12, post = 5.21, gain = 1.09) versus novices (pre = 3.02, post = 4.89, gain = 1.87). Experts gain less from foundational literacy but improve meaningfully [27]—they need advanced, less guided support to avoid plateaus. Subgroup analyses back the moderation hypotheses. Discipline moderated AI literacy and self-efficacy gains; the adaptive design offset variance on primary acceptance constructs.

5.6 Qualitative Triangulation

Thematic analysis of post-test open-ended responses used Braun and Clarke's six-phase framework [37]. Two researchers coded independently (84% inter-coder agreement, remaining settled by consensus). Four themes emerged.

Theme 1—scaffolded confidence—explained how AAIIF coached faculty through building confidence. One humanities participant described their experience moving from "paralyzed uncertainty to systematic exploration." By diagnosing their context, they could "identify specific course elements where AI could add value." This theme aligns with the large PPSE effect ($r = 0.72$) and [52] cognitive-load reducing affordance of AAIIF's step-wise approach. Faculty from several STEM departments reported that iteration with correctness checks "made the evaluation of AI-generated outputs feel disciplined rather than guess-work." As a result, they became comfortable pursuing ideas that may have initially felt daunting. Instead of thinking 'I could get this wrong', faculty thought 'let me see what happens.'

Theme 2—disciplinary calibration—highlighted how participants contextualized AI adoption for their disciplines. Social science instructors discussed how the step emphasizing pedagogical alignment prompted them to identify learning objectives before settling on tech-nology, allowed them to avoid being "gadget happy," and synced with TPACK's pedagogy-before-technology beliefs [31]. Humanities instructors valued the ethics consideration step the most -one second language education faculty noted: "ethical deliberation modules validated my gut feelings about double-checking AItranslated text for cultural and content accuracy, something I wouldn't have thought to do if not for the framework." This was understandable as humanities participants had the greatest increase in AI literacy. Engineering participants noted verification steps that enabled them to have increased "traceability of AI-generated code and circuit-design solutions" [29].

Theme 3—peer-mediated reflection—foregrounded peer learning groups. Sharing successes and challenges "transformed AI integration from an individual burden into a collective inquiry." One mid-career participant described a peer-meeting turning point—"When I heard a colleague describe how they successfully used AI to generate discussion prompts that stimulated critical thinking, I realized the barrier was not the technology but my own reluctance." The theme connects to Bandura's [42] vicarious-experience account and explains the social influence effect ($d = 0.69$). Several early-career faculty said peers gave them "the permission structure to experiment without fear of judgment."

Theme 4—reflective reconstruction—traced shifts in professional identity. Participants took up a stance of "critical partnership" rather than replacement. A senior participant: "Before this program, I viewed AI as a threat to my authority as an educator." The narrative mirrors the large AI literacy effect ($d = 1.16$)—ethical deliberation moved faculty past instrumental adoption toward metacognitive engagement. Late-career participants, the same subset with smaller behavioral-intention gains, produced the most thorough reflective answers—slower uptake paired with deeper philosophical engagement [35].

A minor sub-theme also surfaced: initial aversion to structure. Four participants voiced frustration with AAIIF's prescriptive nature; one called it “too rigid for creative teaching contexts.” Three later noted “the structure eventually liberated [them] from unstructured anxiety” after the first two stages. Faculty internalize AAIIF's logic before adapting it.

6 Discussion

Why did AAIIF work? The process-oriented design drives the gains. The most striking finding was how cycle adherence to the Plan-Generate-Verify-Adapt loop tracked AI-integrated output quality—the iterative loop itself, not any single component, is the active ingredient. That pushes back on technology acceptance work that centres individual variables and skips how faculty engage with AI [7]. Every phase cascaded into the next, as would be expected given accounts of teacher-cognition which conceptualize expertise development as a reinforcing cycle [31]. Magnitudes of effect were varied; some con-structs such as habitual automaticity may require lengthier interventions. Across disciplines - STEM, social sciences, and humanities - AAIIF operated successfully, and disciplinary differences were limited to AI literacy and self-efficacy. Concerning language/literature departments, the ethics component can emphasize cross-cultural verification and language integrity - open-ended responses connected these specifically to gains in literacy knowledge, particularly for instructors of second languages [28]. For engineering programs, the verification stage needs accuracy checklists and ethics scenarios addressing traceability (circuit-design problems were a recurrent example) [29]. Early-career faculty take up AI faster; senior colleagues build deeper self-efficacy. The case is for differentiated rollouts—a lighter version for experienced faculty to avoid ceilings, intensive mentoring for early-career faculty [35].

Three methodological points qualify the results. We chose a single-group quasi-experimental design because withholding AI access wasn't ethical—leaving room for pre/post change driven by media coverage, peer learning, or maturation [45]. Without randomisation, causal attribution isn't definitive; effect sizes likely sit at the upper bound. The cohort came from one Romanian public university, so cultural and infrastructural setting limits transfer. Sixty-two participants give power for large effects but not small-to-medium discipline-by-time interactions—non-significant interactions are possibly underpowered, not confirmed uniformity. Limitations and social desirability bias are introduced by the use of self-report measures, however this was mitigated somewhat by triangulation between quantitative and qualitative data. Many of these limitations should be addressed in future studies. Ideally AAIIF should be tested against a control condition such as unstructured exploration in multisite randomised trials. Future investigations might also compare it against solely workshop-based formats and/or peer-to-peer mentoring with less formal structure in order to tease out the value added by adhering to the cycle. Six- to twelve-month follow-ups would reveal if improvements persist in actual classroom use and if the habit strengthens with repetition. Researchers should also test AAIIF in pre-service teacher education, with graduate teaching assistants, and in K-12 development. Senior faculty produced a split pattern—high

literacy but moderated intention. Longitudinal diary or think-aloud protocols could surface the deliberation behind senior uptake decisions and concerns about identity or institutional resistance.

7 Limitations

Primarly, the regression to the mean is a concern – the samples where recruited on the basis of having low levels of baseline AI knowledge/experience, and so those faculty with lower baseline scores are likely to have benefitted from this phenomenon. Finally, as a single site evaluation our findings may not be generalisable to centres with different populations or resources. Twelve weeks sufficed for behavioral intention, AI literacy, and self-efficacy—not the full uptake arc. Habit produced the smallest effect ($r = 0.61$); deep consolidation needs practice past the program window. We collected no follow-up data—whether gains persist once peer groups and mentoring end is open.

Three measurement issues warrant attention. AAIIF-FQ has acceptable consistency (alpha 0.71 to 0.92) but is fully self-report—participants may have inflated post-test ratings to appear competent, especially since the instrument ran inside the same institution. The single-item prior-AI-experience measure misses the multi-dimensional nature of AI literacy. The 62 completers were self-selected volunteers, more motivated toward AI integration than the broader faculty—selection bias limits representativeness. Faculty who declined may form a more resistant subgroup. The cohort skewed early-to-mid career (mean 11.8 years; 70% within that band); the late-career subgroup was too small to test interactions with discipline or prior AI experience. Cross-cultural comparisons remain out of reach—the Romanian system may produce different response patterns than institutions in North America or Asia. The analytic approach used Bonferroni correction and non-parametric alternatives but has limits. Correcting protects against Type I error yet reduces power, so smaller effects may go undetected in modest-cell subgroup analyses. The MANCOVA doesn't account for unmeasured confounders—department, prior development, baseline tech attitudes. The qualitative coders knew the AAIIF design—a confirmation-bias risk. Future qualitative work should use blinded coders. UTAUT-2 [7] was built in organisational and consumer contexts—its constructs may not fit higher education faculty. Hedonic motivation is a weak driver for faculty whose identity centres on disciplinary expertise. Social influence misses academic-department politics—tenure, promotion, and disciplinary standing shape uptake in ways the items can't reach. Early career faculty showed more growth on the attitude of intention than their more advanced colleagues, who showed greater improvement on literacy outcomes.- institutions should consider pathways tailored to career stage rather than one-size fits all implementations. Generative AI is evolving quickly-no research paper will keep pace. Across our twelve weeks, major platforms expanded features and institutional rules shifted. AAIIF was tested at a single point in time; ethical reasoning will need revision as AI grows more capable.

8 Conclusion

The five-component cycle produced significant gains across all constructs, with largest effects on AI literacy and pedagogical self-efficacy. A process-oriented model tuned for current AI lifts faculty readiness past unstructured exploration. AAIIIF is not meant to be the solution, but rather a demonstration of the utility of organised but flexible processes for faculty to keep up with the ever-changing pedagogical landscape of generative AI.

9 Appendices

9.1 AAIIF-Faculty Questionnaire (AAIIF-FQ) Full Item Set

This appendix presents the full set of items for the AAIIF-Faculty Questionnaire, which is the thirty-eight-item adapted UTAUT-2 instrument employed as the primary measurement tool in this investigation. The items are organized by construct as defined and operationalized in Section 5.3. All items were presented on a seven-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). Effort Expectancy (EE; 4 items) EE1: Learning to operate AI tools for teaching would be easy for me. EE2: My interaction with AI tools in the classroom would be clear and understandable. I perceive AI tools as straightforward to operate for educational objectives. I could readily attain proficiency in applying artificial intelligence to my instruction. Social Influence (SI; 3 items) SI1: Individuals who shape my instructional practices believe I ought to employ AI tools. Individuals who hold significance for me within my department endorse my employment of AI for instructional purposes. The administration of my institution promotes the adoption of AI tools for pedagogical purposes. Facilitating Conditions (FC; 4 items) FC1: I have access to the necessary re-sources to use AI tools for my teaching. FC2: I know enough to use AI tools for my courses. FC3: AI tools are compati-ble with how I teach. When needed, I can get help from peers with AI tools for my teaching practice. Hedonic Motivation (HM; 3 items) HM1: Using AI tools in my teaching is enjoyable. HM2 finds the process of embedding AI into their courses to be intrinsically motivating. HM3: Experimenting with AI tools for pedagogical purposes is useful. Habit (HT; 3 items) HT1: The employment of AI tools has become a customary part of my teaching practice. Frequency was assessed on a five-point scale: 1 (never), 2 (rarely), 3 (sometimes), 4 (often), 5 (very frequently). Depth of adoption was evaluated on a four-point scale derived from the SAMR model [32]: 1 (substitution: AI supplants a conventional instrument with no functional change), 2 (augmentation: AI supplants a conventional instrument with functional improvement), 3 (modification: AI permits major task redesign), 4 (redefinition: AI permits the development of novel tasks previously unattainable).

9.2 Open-Ended Reflection Prompts

Three questions utilizing open-ended responses were also administered at post-test in order to triangulate data. Think back to when you used the AAIIF to embed AI into your instruction. 1. Which resource or step was most helpful to you? Why? 2. What challenges, if any, did you encounter throughout the intervention? How did you overcome them? Was there anything about embedding artificial intelligence technology into your classroom that you felt the AAIIF did not support you with? 3. Lastly, In what ways, if any has your perspective on embedding artificial intelligence into learning environments changed throughout this class? Has your identity as a teacher changed with respect to AI?

Acknowledgments. We thanks to our colleagues and collaborators who supported the design and execution of this study.

Disclosure of Interests. The authors declare that they have no competing financial or non-financial interests that are relevant to the content of this article.

References

1. V Venkatesh & J Thong (2016) Unified theory of acceptance and use of technology: A synthesis and the road ahead. Journal of the Association for Information Systems. DOI: 10.17705/1jais.00428
2. FD Davis (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly. DOI: 10.2307/249008
3. F Shwedeh (2024) The integration of artificial intelligence (AI) into decision support systems within higher education institutions. Nanotechnology Perceptions.
4. RM Fernandes, VMN Nagata, W Leal Filho, et al. (2026) Proposal of a framework to enhance the adoption of artificial intelligence in higher education considering sustainable development: a hybrid Lawshe-TOPSIS, International Journal of Educational Management. DOI: 10.1108/IJEM-08-2024-0441
5. J Feng, B Yu, WH Tan, Z Dai & Z Li (2025) Key factors influencing educational technology adoption in higher education: A systematic review. PLOS Digital Health. DOI: 10.1371/journal.pdig.0000764
6. E Ifinedo & M Kankaanranta (2021) Understanding the influence of context in technology integration from teacher educators' perspective. Technology, Pedagogy and education. DOI: 10.1080/1475939X.2020.1867231
7. M Marikyan & P Papagiannidis (2021) Unified theory of acceptance and use of technology. TheoryHub book.
8. S Mandal (2024) Roadmap for Humanities and Social Sciences in STEM Higher Education: An Introduction. Roadmap for Humanities and Social Sciences in STEM Higher Education. DOI: 10.1007/978-981-97-4275-2_1
9. PS Aithal & AK Maiya (2023) Innovations in higher education industry–Shaping the future. International Journal of Case Studies in Business, IT, and Education. DOI: 10.47992/ijcsbe.2581.6942.0321

10. PC Cormas, G Gould, L Nicholson, KC Fredrick, et al. (2021) A professional development framework for higher education science faculty that improves student learning. BioScience. DOI: 10.1093/biosci/biab050
11. C Yao, A Follmer Greenhoot, K Mack, C Myrick, et al. (2023) Humanizing STEM education: An ecological systems framework for educating the whole student. Frontiers in Education. DOI: 10.3389/feduc.2023.1175871
12. N Emery, JM Maher & D Ebert-May (2019) Studying professional development as part of the complex ecosystem of STEM higher education. Innovative Higher Education. DOI: 10.1007/s10755-019-09475-9
13. C Klein, J Lester, H Rangwala & A Johri (2019) Learning analytics tools in higher education: Adoption at the intersection of institutional commitment and individual action. The Review of Higher Education. DOI: 10.1353/rhe.2019.0007
14. M Sedwal (2024) School education in creating a sustainable world: Role of humanities and social science discipline as a catalyst in the light of indian national education policy 2020. Roadmap for Humanities and Social Sciences in STEM Higher Education, pp. 107-130. DOI: 10.1007/978-981-97-4275-2_7
15. JP Ruiz-Fuentes, et al. (2026) Artificial intelligence and collective intelligence in digital higher education: a quasi-experimental study using the Kampal platform. Revista Iberoamericana de Educación a Distancia. DOI: 10.5944/ried.45560
16. L Pertuz Altamiranda, A Domínguez, et al. (2025) The Impact of the Flipped Classroom Approach with Adaptive Learning on Student Competency in Higher Education: A Quasi-Experimental Study in a Developing, The International Journal of Technologies in Learning, 32(1), pp. 265-279. DOI: 10.18848/2327-0144/CGP/v32i01/265-279
17. BTE Hessen, OA Hewary, et al. (2026) When Artificial Intelligence Meets Blended Learning: A Quasi-Experimental Study of Physics Achievement at the University of Khartoum, Sudan. Tarbawi: Jurnal Keilmuan Manajemen Pendidikan. DOI: 10.32678/tarbawi.v12i01.11360
18. ІВ Леонтьєва (2025) Analytical centers of universities as drivers of innovative development of the ecosystem of higher teacher education. Pedagogical Education: Theory and Practice.
19. M Muhammad, Z Nizam, et al. (2025) Personalized Learning through ChatGPT: A Quasi-Experimental Study on Adaptive Curriculum in High School Settings. Al-Hijr: Journal of Adulearn World, 4(2), pp. 72-86.
20. V Venkatesh, JYL Thong & X Xu (2012) Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology1. MIS quarterly. DOI: 10.2307/41410412
21. K Holstein, V Aleven & N Rummel (2020) A conceptual framework for human–AI hybrid adaptivity in education. In International Conference on Artificial Intelligence in Education. DOI: 10.1007/978-3-030-52237-7_20
22. TL Kakinda, MB Mulumba, FM Masaazi & E Masembe (2026) From Principles to Practice: The Plan-Generate-Verify-Adapt (PGVA) Framework for AI Integration in Luganda Language Instruction. International Journal for Multidisciplinary Research, 7(6). DOI: 10.36948/ijfmr.2025.v07i06.62106.
23. L Konjarski & C Marc (2026) Reframing Academic Practice: The Role of Professional Learning in the VU Block Model. Handbook of Research on the Design, Implementation, and Evaluation of Professional Learning for Higher Education. DOI: 10.4018/979-8-3373-3638-1.ch006
24. S Krishnan (2025) Ethics-Integrated AI Pedagogy and Metacognitive Growth: A Longitudinal Quasi-Experimental Study in Indian Higher Education. SSRN preprint, DOI: 10.2139/ssrn.5526736.

25. G De la Cruz Martínez, YJP Martínez, et al. (2025) Integration Of Generative Ai For Personalized Teacher Support In Instructional Design, ICERI. DOI: 10.21125/iceri.2025.1964
26. J Lave & E Wenger (1991) Situated learning: Legitimate peripheral participation. books.google.com.
27. C Wang & M Zhu (2026) Bridging AI Literacy and AI Pedagogy in Teacher Education and Professional Development: A Multiple Case Study of AI in Education (AIEd) Certificate Programs. Journal of Technology and Teacher Education.
28. M Özdere (2023) The integration of artificial intelligence in English education: Opportunities and challenges. Language Education and Technology Journal.
29. M Usher & M Barak (2024) Unpacking the role of AI ethics online education for science and engineering students. International Journal of STEM Education. DOI: 10.1186/s40594-024-00493-4
30. R Balkhyoor & E Polyak (2024) Machines And Imagination: Navigating The Challenges Of Generative Ai In Design Education, In ICERI2024 Proceedings. DOI: 10.21125/iceri.2024.1282
31. P Mishra & MJ Koehler (2006) Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers college record. DOI: 10.1111/j.1467-9620.2006.00684.x
32. S Boateng & G Kalonde (2024) Exploring the synergy of the SAMR (substitution, augmentation, modification, and redefinition) model and technology integration in education, Proceedings of the International Conference on Advanced Research in Education, Teaching, and Learning, 1(1), pp. 37-46. DOI: 10.33422/aretl.v1i1.185
33. DR Russell (2013) Activity theory and its implications for writing instruction. Reconceiving Writing, Rethinking Writing Instruction, DOI: 10.4324/9780203811948-5
34. DA Schön (2017) The reflective practitioner: How professionals think in action. taylorfrancis.com. DOI: 10.4324/9781315237473
35. A Uzorka, S Namara & AO Olaniyan (2023) Modern technology adoption and professional development of lecturers. Education and Information Technologies. DOI: 10.1007/s10639-023-11790-w
36. RK Hambleton (1996) Guidelines for Adapting Educational and Psychological Tests. ERIC.
37. V Braun & V Clarke (2006) Using thematic analysis in psychology. Qualitative research in psychology. DOI: 10.1191/1478088706qp063oa
38. BG Tabachnick, LS Fidell & JB Ullman (2007) Using multivariate statistics. pearsonhighered.com.
39. CL Blalock, MJ Lichtenstein, S Owen, et al. (2008) In pursuit of validity: A comprehensive review of science attitude instruments 1935–2005. International Journal of Science Education. DOI: 10.1080/09500690701344578
40. LJ Cronbach (1951) Coefficient alpha and the internal structure of tests. psychometrika. DOI: 10.1007/BF02310555
41. SK Morsbach & RJ Prinz (2006) Understanding and improving the validity of self-report of parenting. Clinical child and family psychology review. DOI: 10.1007/s10567-006-0001-5
42. Bandura, A. (1997) Self-efficacy: The exercise of control. New York: Freeman. (Reviewed by Locke, E.A. in Personnel Psychology, 50(3), pp. 801-804)
43. RP DeShon, RE Ployhart, et al. (1998) The estimation of reliability in longitudinal models. International Journal of Behavioral Development. DOI: 10.1080/016502598384243
44. CJ Durham, LD McGrath, et al. (2002) The effects of repeated administrations on self-report and parent-report scales. Journal of Psychoeducational Assessment. DOI: 10.1177/073428290202000302

45. AL Edwards (1985) Multiple regression and the analysis of variance and covariance. psycnet.apa.org.
46. Siegel, S. (1956) Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill.
47. J Cohen (2013) Statistical power analysis for the behavioral sciences. api.taylorfrancis.com. DOI: 10.4324/9780203771587
48. Z Zaremohzzabieh, S Roslan, Z Mohamad, IA Ismail, et al. (2022) Influencing factors in MOOCs adoption in higher education: A meta-analytic path analysis. Sustainability. DOI: 10.3390/su14148268
49. MS Hagger, K Hamilton, DJ Phipps, et al. (2023) Effects of habit and intention on behavior: Meta-analysis and test of key moderators. Motivation Science. DOI: 10.1037/mot0000294
50. S Porkodi & BKH Tabash (2024) A comprehensive meta-analysis of blended learning adoption and technological acceptance in higher education. International Journal of Modern Education and Computer Science (IJMECS), 16(1), pp. 47-71. DOI: 10.5815/ijmecs.2024.01.05
51. MA Kraft (2020) Interpreting effect sizes of education interventions. Educational researcher. DOI: 10.3102/0013189X20912798
52. J Cordero, J Torres-Zambrano & A Cordero-Castillo (2024) Integration of generative artificial intelligence in higher education: Best practices. Education Sciences, 15(1):32. DOI: 10.3390/educsci15010032.