Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

When Does Generative AI Improve Professional Performance? Task–Technology Fit as the Adoption Driver in an Integrated TTF–UTAUT Model

Mohamed Ramy Faisal

M.Faisal1015@student.aast.edu

Arab Academy for Science, Technology, and Maritime Transport

Cherine Soliman

Arab Academy for Science, Technology, and Maritime Transport

Research Article

Keywords: Generative AI, Task–technology fit, UTAUT, Professional performance, AI literacy, PLS-SEM

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11586553/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

When Does Generative AI Improve Professional Performance? Task–Technology Fit as the Adoption Driver in an Integrated TTF–UTAUT Model

Abstract

Generative artificial intelligence (GenAI) is widely expected to raise professional productivity, yet evidence on when it does so remains thin, and adoption is often treated as sufficient for impact. This study integrates Task–Technology Fit (TTF) with the Unified Theory of Acceptance and Use of Technology (UTAUT) and tests a model in which fit shapes the UTAUT expectancies, which drive behavioral intention, actual use and operational performance, with AI literacy, generation and formal training examined as moderators. Survey data from 167 professionals in 13 industry categories in Egypt were analyzed with partial least squares structural equation modeling (PLS-SEM, 5,000 bootstrap subsamples), regression and moderation analysis. Ten of eleven structural paths were supported. TTF strongly predicted performance expectancy ($\beta = .62$) and effort expectancy ($\beta = .50$); performance expectancy was the dominant driver of intention ($\beta = .57$), and actual use predicted operational performance ($\beta = .47$). The effect of fit on intention ran through expected usefulness (indirect $\beta = .36$, $p < .001$). Effort expectancy did not predict intention ($\beta = -.04$, $p = .555$), a shared-variance effect rather than a ceiling. AI literacy was the only significant moderator, strengthening the conversion of intention into performance (conditional effect rising from .58 to .89), whereas generation and training were not. The model explained 72.4% of the variance in intention. The findings reframe the productivity question from whether GenAI helps to when it does: fit and capability, not adoption or age, decide whether GenAI use becomes professional performance.

Keywords: Generative AI; Task–technology fit; UTAUT; Professional performance; AI literacy; PLS-SEM

1 Introduction

Generative artificial intelligence (GenAI) has entered professional work with exceptional speed, accompanied by strong claims about productivity. Experimental and field evidence supports part of that claim: GenAI shortened writing tasks and improved their quality for college-educated professionals [1], raised the productivity of customer-support agents, especially less experienced ones [2], and supported documentation, coding and communication across knowledge work [3, 4]. Yet the same evidence shows that gains are uneven. In a field experiment with management consultants, GPT-4 improved speed and quality for tasks inside the model's capability frontier but made performance worse for tasks just outside it [5]. Adoption, in other words, is not sufficient for impact; what matters is whether the tool fits the task.

Two literatures address parts of this problem but rarely meet. Technology-acceptance research, and the Unified Theory of Acceptance and Use of Technology (UTAUT) in particular, explains why people intend to use and do use a technology [6], but says little about whether use improves performance. Task–Technology Fit (TTF) theory explains when technology improves individual performance, namely when its functionality matches task requirements [7], but says little about the beliefs that drive adoption. Recent GenAI studies have begun to combine the two [8–11], but they typically treat fit as one more predictor of intention alongside the UTAUT expectancies, study students or single sectors, and seldom follow the chain from fit through adoption to performance among practicing professionals.

A second gap concerns who benefits. A popular account holds that younger, "digital native" professionals adapt faster and gain more from new technology [12, 13]. A competing account locates the advantage in capability: the ability to judge when output is reliable, to

steer the tool and to choose tasks it handles well, which the jagged-frontier evidence makes decisive [5]. The two accounts are rarely tested against each other.

This study addresses both gaps. It integrates TTF with UTAUT in a model in which fit is an antecedent of the UTAUT expectancies, and follows the chain through behavioral intention and actual use to operational performance. It then tests AI literacy, generation and formal training as rival moderators of the conversion of intention into performance. The model is estimated on survey data from 167 professionals in 13 industry categories in Egypt, collected after a qualitative phase of 12 interviews that informed the instrument.

The study makes four contributions. First, it specifies the TTF–UTAUT integration mechanistically: fit shapes how useful and how easy professionals expect GenAI to be, and its effect on intention runs through expected usefulness rather than directly. Second, it shows that effort expectancy has no unique effect on intention among practicing professionals, and that this reflects shared variance with usefulness rather than a ceiling effect. Third, it finds that AI literacy, not generation or training, conditions how strongly intention converts into performance, which contradicts the digital-native account. Fourth, it shows that actual use simultaneously raises performance, AI literacy and perceived risk, so that benefit and exposure grow together. The paper proceeds with the theoretical background and hypotheses, the method, the results and a discussion of implications and limitations.

2 Theoretical Background and Hypotheses

2.1 Generative AI and professional performance: the jagged frontier

GenAI tools such as ChatGPT, Copilot and Gemini are built on large language models that generate fluent text, code and analysis from probabilistic patterns learned from training data. This makes them strong at structured language generation, summarization, drafting and ideation support, and weak where tasks require contextual judgment, domain-specific nuance

or causal reasoning. Dell'Acqua et al. [5] describe the resulting boundary as a "jagged technological frontier": tasks of similar apparent difficulty fall on different sides of it, so that the same professional can gain substantially on one task and lose accuracy on the next. The frontier explains why the literature reports both productivity gains and harms such as over-reliance, hallucinated output and reduced originality [3]: both occur, depending on whether the task lies inside or outside the tool's capability. For professional performance, the implication is that frequency of use alone cannot predict outcomes; the alignment between task and tool must be modeled explicitly.

2.2 Integrating Task–Technology Fit with UTAUT

TTF theory holds that information technology improves individual performance to the extent that its functionality corresponds to the demands of the task, with fit arising from the interaction of task and technology characteristics [7]. UTAUT explains adoption through performance expectancy, effort expectancy, social influence and facilitating conditions, with behavioral intention and facilitating conditions predicting use [6]. The two theories are complementary: UTAUT explains why professionals adopt a technology but not why they come to see it as useful or easy, while TTF explains when the technology helps but not how that judgment becomes adoption.

GenAI research has started to combine them. In a study of 470 information-technology employees, performance expectancy, effort expectancy, social influence, TTF, trust, optimism and innovativeness influenced adoption intention, which in turn predicted actual use [8]. TTF has also been linked to academic performance with GenAI [9], to students' intention to use it [10], to faculty performance together with organizational support [11], to organizational performance from collaborative AI [14] and to AI appropriation in project management [15]. Most of these studies, however, place fit beside the UTAUT expectancies as a parallel predictor. The present model places it before them, because the jagged frontier

implies that professionals first judge whether a tool fits the task and only then form expectations about its usefulness and ease.

2.3 Hypotheses for the adoption chain

A tool whose capabilities match the task delivers useful output with less correction, so professionals who perceive high fit should expect it to improve their performance and to require less effort.

H1a: *Task–technology fit positively influences performance expectancy.*

H1b: *Task–technology fit positively influences effort expectancy.*

UTAUT predicts that performance expectancy, effort expectancy and social influence shape behavioral intention [6]. The original theory also found that the effect of effort expectancy weakens as users gain experience, which suggests that it may be weaker than usefulness among practicing professionals who already use these tools.

H2a: *Performance expectancy positively influences behavioral intention.*

H2b: *Effort expectancy positively influences behavioral intention.*

H2c: *Social influence positively influences behavioral intention.*

In UTAUT, intention and facilitating conditions, the resources and support available for use, jointly predict use behavior [6]. Following TTF and the experimental evidence [1, 7], actual use should in turn improve operational performance: productivity, quality, time efficiency and decision effectiveness.

H3a: *Behavioral intention positively influences actual use.*

H3b: *Facilitating conditions positively influence actual use.*

H4: *Actual use positively influences operational performance.*

If fit operates upstream, its influence on intention should run through expected usefulness rather than directly.

H5: *Performance expectancy mediates the effect of task–technology fit on behavioral intention.*

Three further paths are estimated as exploratory rather than hypothesized. A direct path from facilitating conditions to intention allows for the possibility, raised in the qualitative phase, that workplace support and rules shape intention as well as use. Paths from actual use to AI literacy and to perceived risk reflect the double-edged character of GenAI: use can build capability while increasing exposure. Studies of cognitive offloading, the use of external tools to reduce internal cognitive demand [16], show that GenAI can raise performance through self-efficacy while lowering it through reduced independent thinking [17], reduce psychological ownership [18], increase dependence [19] and create conflict as well as creativity [20, 21].

2.4 Moderators: capability or generation?

The jagged frontier makes capability decisive: professionals who understand a tool's limits, can recognize when an answer is likely wrong and can write effective prompts are better placed to keep their use inside the frontier and to convert intention into results. The same evidence suggests that individual differences shape early adoption [22]. AI literacy should therefore strengthen the relationship between intention and performance.

H6: *AI literacy strengthens the relationship between behavioral intention and operational performance.*

The digital-native account predicts the opposite emphasis: younger cohorts, socialized with digital technology, should adapt faster and gain more [12, 13]. Older professionals, however, often bring the domain expertise needed to validate output, which the frontier evidence also rewards. Formal training is a third candidate, since it is the usual organizational response to new tools.

H7: *The relationship between behavioral intention and operational performance is stronger for younger generational cohorts.*

H8: *Formal GenAI training strengthens the relationship between behavioral intention and operational performance.*

Fig. 1 Research model. Solid arrows are hypothesized paths; dashed boxes are exploratory outcomes of actual use. Moderators H6–H8 are tested on the conversion of intention into operational performance.

3 Method

3.1 Research design and context

The study is the quantitative phase of an exploratory sequential mixed-methods project. A preceding qualitative phase of 12 semi-structured interviews with professionals across industries and generations identified the constructs and mechanisms carried into the survey, including verification behavior, AI literacy and perceived risk. The interviews also surfaced the tensions tested here: performance expectancy was the near-universal driver of intention, effort was described as trivial for simple tasks but heavy for complex ones, and participants were divided on whether generation or skill determines who benefits. Egypt provides a useful setting: professionals there adopt general-purpose GenAI tools rapidly, often ahead of organizational policy, across sectors that differ sharply in regulation, from information technology to banking and healthcare.

3.2 Instrument and measures

All constructs were measured with five-point Likert items (1 = strongly disagree, 5 = strongly agree). Performance expectancy (four items), effort expectancy (four), social influence (two), facilitating conditions (three) and behavioral intention (three) were adapted from Venkatesh et al. [6]. Task–technology fit was adapted from Goodhue and Thompson [7]; of six items, three (TTF1, TTF2 and TTF4) were retained in the measurement model.

Operational performance (five items) captured perceived gains in productivity, quality, decision-making and time, drawing on Goodhue and Thompson [7], Noy and Zhang [1] and the interview themes. AI literacy (three items) and perceived risk (nine items covering dependency, deskilling, cost, provider dependence, job replacement, confidentiality, governance and high-stakes use) were developed for this study from the qualitative phase. Actual use was measured with a single behavioral indicator ("I use Generative AI across a wide range of my work tasks"). The items and measurement statistics for each construct are reported with the measurement model in Section 4.2. Respondents also reported their generational cohort, formal GenAI training, length of GenAI use, frequency of use and industry.

A pilot with 63 respondents confirmed the reliability of the instrument before the main survey: behavioral intention (α = .95), operational performance (α = .93), AI literacy (α = .88), performance expectancy (α = .87), effort expectancy (α = .81), perceived risk (α = .77) and facilitating conditions (α = .76).

3.3 Sample and data collection

The questionnaire was distributed online through Google Forms, LinkedIn and professional networks, using professional-network and convenience sampling. The final sample comprised 167 professionals (Table 1). All four generational cohorts were represented, with Millennials the largest group; one third had received formal GenAI training. Respondents came from 13 industry categories, led by information technology, marketing, communication and translation, and healthcare. Half used GenAI daily or more often. The sample exceeds the 10-times rule for PLS-SEM, since the construct with the most incoming paths (behavioral intention) has four predictors [23]; findings are interpreted as bounded to comparable professional populations rather than as population estimates.

Table 1 Sample profile (N = 167)

Characteristic	Category	n	%
Generational cohort	Generation Z	51	30.5
	Millennials	65	38.9
	Generation X	47	28.1
	Baby Boomers	4	2.4
Formal GenAI training	Yes	56	33.5
	No	111	66.5
Length of GenAI use	Less than 6 months	26	15.6
	6 months to 1 year	38	22.8
	1 to 2 years	57	34.1
	2 to 3 years	30	18.0
	More than 3 years	16	9.6
Industry (13 categories)	IT/Technology	33	19.8
	Marketing, communication and translation	29	17.4
	Healthcare/Medical	24	14.4
	Engineering	11	6.6
	Finance/Banking	10	6.0
	Manufacturing/Operations	10	6.0
	Consulting/Advisory	8	4.8
	Education	7	4.2
	Other categories (5)	35	21.0

Note: “Other categories” groups the five smallest industry categories, including respondents who selected “Other”.

3.4 Analysis

Analysis proceeded in two stages. In SPSS, the data were screened (skewness −0.73 to −1.24 and kurtosis 0.10 to 2.88, within conventional limits for parametric analysis), scale reliability was assessed, sampling adequacy was tested with the Kaiser–Meyer–Olkin measure and Bartlett’s test [24], and an ordinary least squares regression of intention on its predictors served as a robustness check. Group differences were tested with one-way ANOVA (generation) and independent-samples t-tests (training), and moderation with the PROCESS macro, Model 1, with mean-centered variables [25]. In SmartPLS 4 [26], the measurement model was assessed through outer loadings, Cronbach’s α, composite reliability, average variance extracted (AVE), the Fornell–Larcker criterion and the heterotrait–monotrait ratio (HTMT), and the structural model through path coefficients with 5,000 bootstrap subsamples, R^2, f^2, specific indirect effects and the standardized root mean square residual (SRMR) [23, 27, 28].

4 Results

4.1 Descriptive statistics and correlations

Construct means ranged from 3.56 (task–technology fit) to 3.96 (behavioral intention), indicating broadly positive evaluations (Table 2). All correlations among the model constructs were positive and significant. The strongest linked performance expectancy, behavioral intention and operational performance (r = .81 for each pair), and fit correlated strongly with operational performance (r = .70) and performance expectancy (r = .62). Perceived risk correlated more strongly with AI literacy (r = .52) than with operational performance (r = .37), an early sign that risk awareness accompanies capability.

Table 2 Means, standard deviations and correlations (N = 167)

Construct	M	SD	1	2	3	4	5	6	7	8
1. Task–technology fit	3.56	0.77	—							
2. Performance expectancy	3.86	0.79	.62	—						
3. Effort expectancy	3.86	0.82	.49	.61	—					
4. Social influence	3.64	0.82	.50	.66	.47	—				
5. Facilitating conditions	3.82	0.79	.47	.63	.66	.54	—			
6. Behavioral intention	3.96	0.91	.59	.81	.53	.69	.64	—		
7. Operational performance	3.86	0.88	.70	.81	.55	.63	.60	.81	—	
8. AI literacy	3.90	0.83	.40	.53	.62	.44	.68	.52	.48	—
9. Perceived risk	3.60	0.77	.20*	.45	.43	.46	.46	.44	.37	.52

Note: All correlations significant at $p < .01$ except * $p < .05$.

4.2 Measurement model

Sampling adequacy was excellent (KMO = .935; Bartlett's $\chi^2(351) = 3908.92$, $p < .001$). Outer loadings ranged from .74 to .95 for all constructs except perceived risk (.57 to .80), where two indicators (RISK1 = .57; RISK2 = .62) fell below .70 and were retained for content validity. Cronbach's α ranged from .72 to .94, composite reliability from .87 to .96 and AVE from .50 to .89 (Table 3). Perceived risk sat exactly at the AVE threshold of .50 and is interpreted with caution. Discriminant validity was supported: the Fornell–Larcker criterion was satisfied and all HTMT values were below .90 (highest: performance expectancy and intention, .898; performance expectancy and operational performance, .893).

Table 3 Constructs, items and measurement model statistics

Construct (source)	Items	Loadings	α	CR	AVE
Task–technology fit [7]	TTF1, TTF2, TTF4	.79–.88	.797	.881	.712
Performance expectancy [6]	PE1–PE4	.82–.89	.877	.916	.731
Effort expectancy [6]	EE1–EE4	.74–.93	.899	.930	.770
Social influence [6]	SI1, SI2	.86–.91	.716	.874	.777
Facilitating conditions [6]	FC1–FC3	.77–.87	.780	.873	.697
Behavioral intention [6]	BI1–BI3	.93–.95	.938	.960	.889
Actual use (single indicator)	AUT1	1.00	—	—	—
Operational performance (developed for this study)	OP1–OP5	.88–.91	.938	.953	.802
AI literacy (developed for this study)	AIL1–AIL3	.88–.91	.877	.924	.803
Perceived risk (developed for this study)	RISK1–RISK9	.57–.80	.877	.899	.500

Note: All items used five-point Likert scales. Example items: "People who influence my work think I should use Generative AI" (social influence); "I have the resources (device, internet, subscription) to use Generative AI for work" (facilitating conditions); "I intend to continue using Generative AI in my work" (intention); "I use Generative AI across a wide range of my work tasks" (actual use); "Generative AI has improved my overall work performance" (operational performance); "Using GenAI without clear organizational rules creates real risks" (RISK8). CR = composite reliability (rho_c); AVE = average variance extracted.

4.3 Structural model and mediation

Ten of the eleven structural paths were significant (Table 4; Figure 2). Task–technology fit strongly predicted performance expectancy (β = .624) and effort expectancy (β = .502), supporting H1a and H1b. Performance expectancy (β = .574) and social influence (β = .223) predicted intention, supporting H2a and H2c, while effort expectancy did not (β = −.041, p = .555), so H2b is not supported. Intention (β = .312) and facilitating conditions (β = .223) predicted actual use (H3a, H3b), and actual use predicted operational performance (β = .468; H4). Of the exploratory paths, facilitating conditions also predicted intention (β = .189), and actual use predicted both AI literacy (β = .362) and perceived risk (β = .353).

The model explained 72.4% of the variance in behavioral intention, 38.9% in performance expectancy, 25.2% in effort expectancy, 23.7% in actual use and 21.9% in operational performance. Effect sizes were large for fit on performance expectancy (f^2 = .636) and performance expectancy on intention (f^2 = .480), and the SRMR of .073 was below the .08 threshold. The specific indirect effect of fit on intention through performance expectancy was significant (β = .358, t = 5.33, p < .001), supporting H5. The downstream chain from intention through actual use to performance was also significant (β = .146, p = .023).

An OLS regression of intention on its five predictors reproduced this pattern ($R^2 = .723$; $F(5, 161) = 84.08$, $p < .001$). Performance expectancy ($\beta = .534$), social influence ($\beta = .211$) and facilitating conditions ($\beta = .193$) were significant; fit ($\beta = .097$, $p = .076$) and effort expectancy ($\beta = -.069$, $p = .244$) were not. With variance inflation factors no higher than 2.70, multicollinearity was not a concern. The non-significant direct effect of fit, despite its correlation of .59 with intention, is the signature of the mediation through performance expectancy.

Table 4 Structural model results and hypothesis tests (PLS-SEM, 5,000 bootstrap subsamples)

Hypothesis	Path	β	t	p	Result
H1a	TTF → Performance expectancy	.624	8.33	< .001	Supported
H1b	TTF → Effort expectancy	.502	6.36	< .001	Supported
H2a	Performance expectancy → Intention	.574	8.02	< .001	Supported
H2b	Effort expectancy → Intention	−.041	0.59	.555	Not supported
H2c	Social influence → Intention	.223	3.77	< .001	Supported
H3a	Intention → Actual use	.312	2.91	.004	Supported
H3b	Facilitating conditions → Actual use	.223	2.09	.037	Supported
H4	Actual use → Operational performance	.468	6.64	< .001	Supported
H5	TTF → Performance expectancy → Intention (indirect)	.358	5.33	< .001	Supported
Exploratory	Facilitating conditions → Intention	.189	2.62	.009	Significant
Exploratory	Actual use → AI literacy	.362	4.49	< .001	Significant
Exploratory	Actual use → Perceived risk	.353	5.08	< .001	Significant
Exploratory	Intention → Actual use → Performance (indirect)	.146	—	.023	Significant

Note: $R^2 = .724$ (intention), .389 (performance expectancy), .252 (effort expectancy), .237 (actual use), .219 (operational performance), .131 (AI literacy), .124 (perceived risk). SRMR = .073.

Fig. 2 SmartPLS estimates: outer loadings, path coefficients and R^2 values (N = 167). Redrawn at print resolution from the SmartPLS 4 output with the estimated values unchanged.

4.4 Why effort expectancy did not predict intention

A ceiling effect, in which every respondent finds GenAI easy, would explain the null path, but the data do not show one. Effort expectancy averaged 3.86 (SD = 0.82), only 54% of respondents scored at or above "agree", and item means ranged from 4.03 (finding the tools easy to use) to 3.60 (the tools being simple enough for non-technical users), with 16% disagreeing with the latter. Nor was its distribution compressed: effort and performance expectancy had identical means (3.86), and effort expectancy had, if anything, slightly greater

variance (0.68 vs 0.62), yet performance expectancy was the strongest predictor in the model. Effort expectancy was, on its own, a good correlate of intention (r = .53), but it correlated strongly with performance expectancy (r = .61). When both compete, the stronger and more proximal construct, performance expectancy, absorbs the shared variance. Ease of use is real and valued, but it is not what turns practicing professionals into continuing users.

4.5 Moderation and group differences

AI literacy significantly moderated the relationship between intention and operational performance (interaction $p < .001$; $\Delta R^2 = .049$). The conditional effect of intention on performance rose from .58 at low literacy to .71 at the mean and .89 at high literacy, supporting H6. Generation did not moderate the relationship (p = .069); the non-significant trend ran opposite to the digital-native expectation, with conditional effects of .60, .72 and .85 from Generation Z to Generation X. H7 is not supported. Formal training was not a significant moderator either (p = .159), so H8 is not supported. A supplementary test of self-rated domain expertise as a moderator of the use–performance relationship was also non-significant (p = .641), although expertise had a strong direct effect on performance.

Group comparisons reinforce these results. Generational cohorts did not differ on operational performance (p = .894), intention (p = .614), task–technology fit (p = .257), performance expectancy (p = .877), perceived risk (p = .632) or AI literacy (p = .258); only effort expectancy differed (p = .015), with younger cohorts reporting greater ease. Formally trained respondents reported higher AI literacy (M = 4.13 vs 3.86) but not higher performance (p = .223) or intention (p = .591). Descriptively, experience with the tool mattered where career experience did not: mean operational performance rose from 3.13 among those using GenAI for less than six months to 4.38 among those using it for more than three years, whereas years of professional experience were unrelated to performance (r = .05).

Performance likewise rose with frequency of use, from 3.12 among those using GenAI less than monthly to 4.34 among those using it several times a day.

Table 5 Moderation of the intention → operational performance relationship (PROCESS Model 1)

Hypothesis	Moderator	Interaction p	ΔR^2	Conditional effects (low → mean → high)	Result
H6	AI literacy	< .001	.049	.58 → .71 → .89	Supported
H7	Generation (multicategorical)	.069	—	.60 → .72 → .85 (Gen Z → Gen X)	Not supported
H8	Formal training	.159	—	—	Not supported

Note: Predictors were mean-centered. Supplementary test: domain expertise × actual use, p = .641 (not significant).

5 Discussion

5.1 Fit comes first

The results support a sequenced integration of TTF and UTAUT. Fit was the strongest antecedent in the model, shaping both performance expectancy and effort expectancy, and its influence on intention ran entirely through expected usefulness. This is consistent with how professionals encounter a jagged frontier [5]: they first judge whether GenAI fits a particular task, and that judgment determines how useful they expect it to be. The interviews suggest how this happens: professionals experience fit in episodes (a task delegated well, a sudden failure just outside the tool's competence, a confidentiality block), and these episodes consolidate into a usefulness expectation that then governs intention. Studies that enter fit as a parallel predictor of intention [8, 10] may therefore understate its role, because its effect is carried by the expectancies rather than added to them.

5.2 Usefulness, not ease

Among practicing professionals, effort expectancy had no unique effect on intention. The diagnostics rule out a ceiling and point to construct dominance: ease of use matters and correlates with intention, but usefulness absorbs its predictive contribution. This refines the UTAUT finding that effort expectancy fades with experience [6]. The interviews anticipated

the result, with effort appearing at both poles, trivial for simple tasks and laborious for complex ones, and so canceling as a net driver. For GenAI, ease is largely, though not universally, solved; what drives continued adoption is the expected gain in productivity and quality.

5.3 Capability, not generation

The digital-native account did not survive the test. Generations did not differ in performance, intention, fit, usefulness, literacy or risk, and generation did not moderate the conversion of intention into performance; if anything, the trend favored older cohorts. Training raised literacy but did not, by itself, raise performance. AI literacy was the only significant moderator, and length of experience with the tool, not length of career, tracked performance. Capability with the tool, however acquired, is the active ingredient. The belief in a generational advantage is nonetheless widespread: two-thirds of respondents (66%) agreed that younger colleagues adopt AI faster, and in the interviews most participants assumed a generational effect while others located the real divide between using GenAI and using it properly. The data resolve that split in favor of capability. This aligns with evidence that the benefits of GenAI depend on the user's ability to evaluate output [5, 18] and qualifies claims that younger workers are natural beneficiaries [12, 13].

5.4 Use is double-edged

Actual use predicted operational performance, AI literacy and perceived risk at the same time. Productivity and risk are therefore parallel consequences of adoption rather than opposite ends of one evaluation, consistent with the double-edged effects reported for creativity, dependence and ownership [17, 19, 20]. That risk perception correlated more with literacy than with performance suggests that informed users see more risk, not less: risk awareness is a marker of capability rather than a barrier to adoption.

5.5 Theoretical implications

The study contributes to information systems and technology-adoption research in three ways. It specifies the TTF–UTAUT integration as a sequence in which fit is an antecedent of the expectancies, with a confirmed mediation path, rather than a parallel predictor. It explains the effort-expectancy null for experienced professional users as shared variance with usefulness, which future UTAUT applications to GenAI should anticipate. And it adds AI literacy to the boundary conditions of the adoption–performance link, while showing that generation, the moderator most often assumed, does not condition it.

5.6 Practical implications

Organizations seeking performance gains from GenAI should start from task fit: identify the tasks inside the frontier and direct use there, rather than encouraging use indiscriminately. Adoption campaigns should demonstrate usefulness for specific tasks, because professionals are already past the ease threshold. Capability building should target AI literacy (understanding limits, detecting errors and steering the tool) for all age groups rather than assuming that younger staff need less support; training is valuable only to the extent that it builds that capability, so programs should be evaluated by the literacy they produce rather than by attendance. Finally, because use raises exposure as well as performance, organizations should pair adoption with secure, sanctioned tools and clear rules for confidential data.

5.7 Limitations and future research

Several limitations bound the findings. Performance was self-reported, and all constructs came from a single cross-sectional survey, so common-method variance cannot be ruled out [29]; objective performance data and longitudinal designs would strengthen causal claims. The convenience sample of 167 professionals in one country supports generalization

only to comparable professional populations, and sector cells were too small for industry-level moderation. Perceived risk sat at the AVE threshold (.50) with two weaker indicators, and performance expectancy and intention, though discriminant by conventional criteria, were closely related (HTMT = .898). Actual use was measured with a single indicator. Future research should test the indirect path from effort expectancy through performance expectancy to intention and use hierarchical regression to demonstrate the shared-variance mechanism formally, compare regulated and unregulated sectors with larger samples, and examine creative outcomes, where GenAI may strengthen convergent work while constraining originality.

6 Conclusion

Generative AI improves professional performance conditionally. It does so when the tool fits the task, because fit builds the expectation of usefulness that drives intention, use and measurable gain, and when professionals have the AI literacy to convert intention into results. Ease of use and generation, the two factors most often assumed to matter, do not. For research, the productivity question is no longer whether GenAI helps but under what conditions; for practice, the answer is to invest in fit and capability rather than in adoption alone.

Abbreviations

AI: Artificial intelligence

AU: Actual use

AVE: Average variance extracted

BI: Behavioral intention

CR: Composite reliability

EE: Effort expectancy

FC: Facilitating conditions

GenAI: Generative artificial intelligence

HTMT: Heterotrait–monotrait ratio

KMO: Kaiser–Meyer–Olkin measure

OLS: Ordinary least squares

OP: Operational performance

PE: Performance expectancy

PLS-SEM: Partial least squares structural equation modeling

SI: Social influence

SRMR: Standardized root mean square residual

TTF: Task–technology fit

UTAUT: Unified Theory of Acceptance and Use of Technology

VIF: Variance inflation factor

References

1. Noy S, Zhang W (2023) Experimental evidence on the productivity effects of generative artificial intelligence. Science 381(6654):187–192. https://doi.org/10.1126/science.adh2586

2. Brynjolfsson E, Li D, Raymond L (2025) Generative AI at work. The Quarterly Journal of Economics 140(2):889–942. https://doi.org/10.1093/qje/qjae044

3. Al Naqbi H, Bahroun Z, Ahmed V (2024) Enhancing work productivity through generative artificial intelligence: A comprehensive literature review. Sustainability 16(3):1166. https://doi.org/10.3390/su16031166

4. Coutinho M, Marques L, Santos A, Dahia M, França C, de Souza Santos R (2024) The role of generative AI in software development productivity: A pilot case study. In:

412 Proceedings of the 1st ACM International Conference on AI-Powered Software. ACM,
413 pp 131–138
414 5. Dell'Acqua F, McFowland E, Mollick ER, Lifshitz-Assaf H, Kellogg K, Rajendran S,
415 Krayer L, Candelon F, Lakhani KR (2023) Navigating the jagged technological
416 frontier: Field experimental evidence of the effects of AI on knowledge worker
417 productivity and quality. Harvard Business School Technology & Operations Mgt. Unit
418 Working Paper No. 24-013. Harvard Business School
419 https://doi.org/10.2139/ssrn.4573321
420 6. Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information
421 technology: Toward a unified view. MIS Quarterly 27(3):425–478.
422 https://doi.org/10.2307/30036540
423 7. Goodhue DL, Thompson RL (1995) Task-technology fit and individual performance. MIS
424 Quarterly 19(2):213–236. https://doi.org/10.2307/249689
425 8. Prasad KDV, Singh S, Srinivas V, Kothari H, Pathak A, Shrimali D (2026) Factors
426 affecting the adoption of generative AI tools among information technology employees:
427 A UTAUT3, TTF, and SOR perspective. Journal of Computational and Cognitive
428 Engineering 5(3):576–588. https://doi.org/10.47852/bonviewJCCE52026127
429 9. Al-Emran M, Al-Sharafi MA, Foroughi B, Al-Qaysi N, Mansoor D, Beheshti A, Ali N
430 (2025) Evaluating the influence of generative AI on students' academic performance
431 through the lenses of TPB and TTF using a hybrid SEM-ANN approach. Education and
432 Information Technologies 30(12):17557–17587. https://doi.org/10.1007/s10639-025-
433 13485-w
434 10. Yang Y-J, Wang C-C, Chen Y-H, Chen H-M (2025) Students' behavioral intentions
435 toward generative AI in education: Task-technology fit and moral obligations. Journal
436 of Education and e-Learning Research 12:675–684

11. Aliazas JV, dela Cruz R, Panoy JF, Andrade R (2025) Faculty performance in the age of generative AI: The role of organizational support systems and task-technology fit in higher education. Advances in Consumer Research 2:1960–1970

12. Prensky M (2001) Digital natives, digital immigrants part 1. On the Horizon 9(5):1–6. https://doi.org/10.1108/10748120110424816

13. Berkup SB (2014) Working with generations X and Y in generation Z period: Management of different generations in business life. Mediterranean Journal of Social Sciences 5(19):218–229. https://doi.org/10.5901/mjss.2014.v5n19p218

14. Przegalinska A, Triantoro T, Kovbasiuk A, Ciechanowski L, Freeman RB, Sowa K (2025) Collaborative AI in the workplace: Enhancing organizational performance through resource-based and task-technology fit perspectives. International Journal of Information Management 81:102853

15. Naseer A, Iqbal S, Haq EU (2026) Innovating with AI: How innovative attitude, peer influence, and task-technology fit shape AI appropriation in project management. Inverge Journal of Social Sciences 5:373–389

16. Risko EF, Gilbert SJ (2016) Cognitive offloading. Trends in Cognitive Sciences 20(9):676–688. https://doi.org/10.1016/j.tics.2016.07.002

17. Ma Y, Zhang Z, Liu C (2026) The double-edged sword effect of GenAI assistance on university students' academic performance: Evidence from China. Education and Information Technologies 31:675–700

18. He W (2026) Subjective task load and psychological ownership in generative AI collaborative music creation: Mechanisms shaping creators' state sense of agency. Frontiers in Psychology 17:1835406

19. Mingjie G, Jun M, Huirong T (2025) The double-edged sword effect of generative artificial intelligence usage on employee creativity: A perspective of conservation of resources theory. Current Psychology 44:8485–8497

20. Luo G, Niu C, Lu L, Wu L, Huang L (2025) Empirical study on the double-edged effects of generative artificial intelligence on team creativity among university students. Education and Information Technologies 30:18853–18878

21. Qian P, Xie X (2026) Task–technology fit leads to conflict: The double-edged-sword effect of generative artificial intelligence on scientific creative performance in humanities and social sciences research. International Journal of Human–Computer Interaction 42:1867–1891

22. Kovbasiuk A, Triantoro T, Przegalińska A, Sowa K, Ciechanowski L, Gloor P (2025) The personality profile of early generative AI adopters: A big five perspective. Central European Management Journal 33:252–264

23. Hair JF, Risher JJ, Sarstedt M, Ringle CM (2019) When to use and how to report the results of PLS-SEM. European Business Review 31(1):2–24. https://doi.org/10.1108/EBR-11-2018-0203

24. Kaiser HF (1974) An index of factorial simplicity. Psychometrika 39(1):31–36. https://doi.org/10.1007/BF02291575

25. Hayes AF (2022) Introduction to mediation, moderation, and conditional process analysis: A regression-based approach, 3rd edn. Guilford Press

26. Ringle CM, Wende S, Becker J-M (2024) SmartPLS 4 [Computer software]. SmartPLS https://www.smartpls.com

27. Fornell C, Larcker DF (1981) Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research 18(1):39–50. https://doi.org/10.1177/002224378101800104

485 28. Henseler J, Ringle CM, Sarstedt M (2015) A new criterion for assessing discriminant
486 validity in variance-based structural equation modeling. Journal of the Academy of
487 Marketing Science 43(1):115–135. https://doi.org/10.1007/s11747-014-0403-8
488 29. Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP (2003) Common method biases in
489 behavioral research: A critical review of the literature and recommended remedies.
490 Journal of Applied Psychology 88(5):879–903. https://doi.org/10.1037/0021-
491 9010.88.5.879

Figures

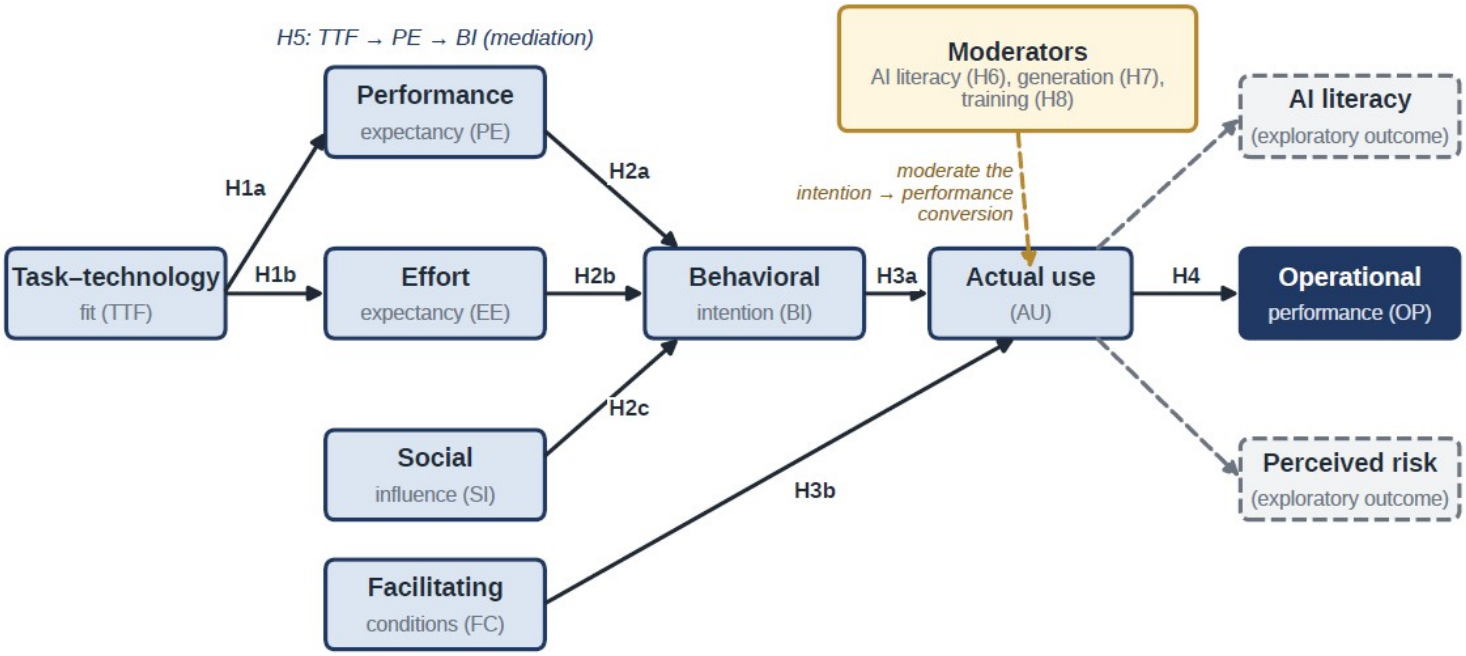


Figure 1

Research model. Solid arrows are hypothesized paths; dashed boxes are exploratory outcomes of actual use. Moderators H6–H8 are tested on the conversion of intention into operational performance.

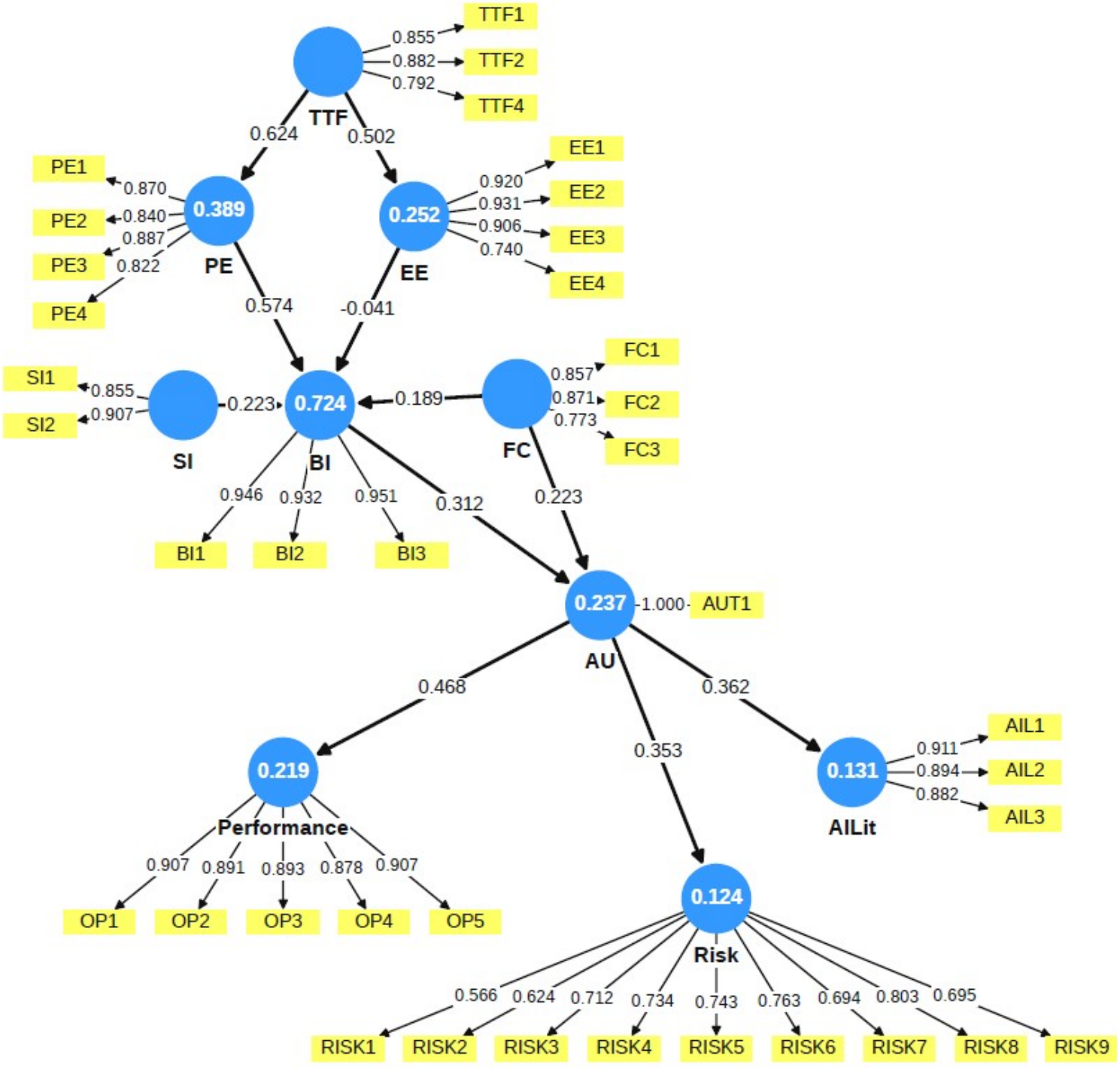


Figure 2

SmartPLS estimates: outer loadings, path coefficients and R² values (N = 167). Redrawn at print resolution from the SmartPLS 4 output with the estimated values unchanged.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ProcessTest.htm

- ModerationProcessTest.htm
- SPSS167.htm