

The Impact of Artificial Intelligence Adoption on Employee Performance: The Mediating Role of Digital Skills

Agare Oghenerume ¹, Mou Napoleon ²

¹ Southern Delta University, Ozoro

² Southern Delta University, Ozoro

Email: ¹ agareo@sdu.edu.ng, ² Mounapoleon7@gmail.com

ARTICLE INFO

Keywords:

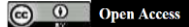
Artificial intelligence adoption, employee performance, digital skills, mediation, PLS-SEM, Monte Carlo simulation, human resource management, common method variance.

Vol. 01, No. 05, 2026

<https://doi.org/10.5281/zenodo.23263212>

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CCBY4.0).

<http://creativecommons.org/licenses/by/4.0/>



ABSTRACT

Organisations are adopting artificial intelligence (AI) rapidly, yet evidence on whether and how adoption improves individual employee performance remains fragmented. Drawing on the technology acceptance tradition, task–technology fit, human capital theory and the job demands–resources model, this study develops and tests a model in which digital skills mediate the relationship between AI adoption and employee performance. The study follows a transparent, simulation-based design. A measurement instrument (17 reflective items) and a population model with a priori specified effects were defined. Survey-type data for 350 employees were generated by Monte Carlo simulation and analysed with partial least squares structural equation modelling (PLS-SEM), consistent PLS (PLSc) and 5,000-sample bootstrapping. Three further Monte Carlo experiments (1,000 to 8,400 replications) examined parameter recovery, statistical power for the indirect effect, and the sensitivity of estimates to common method variance. The measurement model met all conventional criteria ($\alpha = .85-.88$, composite reliability = $.89-.91$, AVE = $.58-.67$, HTMT $\leq .46$). AI adoption was positively related to digital skills ($\beta = .412$, $p < .001$) and digital skills to employee performance ($\beta = .321$, $p < .001$). The indirect effect was significant ($\beta = .132$, 95 % CI $[.088, .191]$) and accounted for 51 % of the total effect, indicating complementary partial mediation. In the Monte Carlo experiments, standard PLS-SEM under-estimated the a and b paths by 12–13 % and the indirect effect by 23 %, whereas PLSc was virtually unbiased. Detecting a small indirect effect with 80 % power required more than 500 respondents. A method factor with loadings of 0.30 inflated the indirect effect by 31 %, yet the first factor still explained only 45 % of the variance on average, so Harman's single-factor test would not have flagged the bias. The study clarifies the mechanism through which AI adoption may translate into performance, namely the development of employees' digital skills. It also provides validated design guidance, including estimator choice, sample size and method-bias control, for empirical research on AI in the workplace. Because the data are simulated, the estimates illustrate the model's implications and the behaviour of the analytical procedure; they are not empirical evidence. The model must be confirmed with field data.

Introduction

Artificial intelligence (AI) has moved from experimental pilots to everyday workplace use. Generative assistants draft text and code, machine-learning systems forecast demand and screen applicants, and analytics platforms embed algorithmic recommendations into routine decisions (Budhwar et al., 2023; Tambe et al., 2019). Employers expect these technologies to reshape skill requirements substantially within a few years, and they identify AI and big data as among the fastest-growing skill areas (World Economic Forum, 2025). For organisations, the central question is no longer whether to adopt AI but how adoption translates into value. Because much of that value is realised through what employees do with AI at work, the question becomes one of human resource management (HRM) as much as of technology (Chowdhury et al., 2023; Makarius et al., 2020).

Recent field and experimental studies report substantial productivity gains from AI assistance. Customer-support agents with access to a generative AI assistant resolved about 15 % more issues per

hour, and the gains were largest for less experienced workers (Brynjolfsson et al., 2025). Professionals using ChatGPT completed writing tasks faster and with higher quality (Noy & Zhang, 2023). Developers using an AI pair programmer completed a coding task considerably faster than a control group (Peng et al., 2023). Yet the same literature shows that the gains are uneven. Consultants performed better with AI on tasks inside the 'jagged frontier' of AI capability but worse on tasks outside it, where uncritical reliance on AI led to errors (Dell'Acqua et al., 2023). Organisational surveys likewise report that many firms have not realised the benefits they expected from AI investments (Chowdhury et al., 2023). Adoption, in other words, is not sufficient for performance. Something about the employee who uses the technology conditions its effect.

This study proposes that digital skills are a central part of that 'something'. Digital skills encompass the technical ability to operate digital tools as well as the information-management, communication, collaboration, creativity, critical-thinking and problem-solving skills needed to use them productively (van Laar et al., 2017). Using AI effectively requires such skills: users must formulate tasks for the system, evaluate and verify its outputs, integrate them with other information, and recognise when the system should not be trusted (Raisch & Krakowski, 2021). At the same time, working with AI is itself a context in which these skills are exercised and developed. Adoption may therefore improve performance partly because it builds the digital capabilities through which employees convert AI into better work.

Despite its intuitive appeal, this mechanism has received limited systematic attention. Technology-acceptance research has concentrated on the antecedents of adoption, such as perceived usefulness, effort expectancy and social influence, rather than on its consequences for performance (Davis, 1989; Venkatesh et al., 2003; Venkatesh, 2022). HRM research on AI has been largely conceptual (Budhwar et al., 2023; Tambe et al., 2019), and economic studies treat skills mainly as a determinant of which tasks are automated (Acemoglu & Restrepo, 2019; Autor, 2015). Empirical work that does test mediating mechanisms typically relies on cross-sectional self-reports analysed with PLS-SEM. The statistical properties of this design under realistic conditions, including measurement error, modest effect sizes and common method variance, are seldom examined, yet they determine how far the published estimates can be trusted (Hair et al., 2019; Podsakoff et al., 2012).

This paper makes three contributions. First, it integrates four theoretical perspectives into a parsimonious mediation model linking AI adoption, digital skills and employee performance, and it derives four hypotheses. Second, it operationalises the model with a 17-item measurement instrument and demonstrates the complete analytical workflow, from measurement-model assessment to bootstrapped mediation testing, on a simulated sample of 350 employees. Third, it uses Monte Carlo experiments to quantify how much standard PLS-SEM attenuates the hypothesised effects, how many respondents are needed to detect the indirect effect, and how strongly common method variance can distort conclusions. Together, these results provide a theoretically grounded and methodologically tested template for future field studies. The data in this paper are simulated, and its findings concern the implications of the model and the performance of the methods. They are not an empirical test of the hypotheses in a real organisation.

2. Literature Review and Hypothesis Development

2.1 AI adoption in the workplace

In this study, AI adoption refers to the extent to which an employee uses AI-based applications in performing work tasks and relies on their outputs. Such applications include machine-learning predictions, natural-language assistants, intelligent automation and recommendation systems. The

construct is defined at the individual level, as the employee's actual use and integration of AI into work, in line with the use-behaviour construct of the Unified Theory of Acceptance and Use of Technology (UTAUT; Venkatesh et al., 2003). Organisational AI investment matters only to the extent that it reaches employees' work. Venkatesh (2022) accordingly argues that research on AI tools should move beyond intention to examine use and its outcomes.

Economic theory distinguishes two effects of new technology on work. Automation substitutes for labour in some tasks. Augmentation complements human labour, raises its productivity and creates new tasks (Acemoglu & Restrepo, 2019; Autor, 2015). Raisch and Krakowski (2021) show that, in management, automation and augmentation are interdependent: AI systems take over routine elements of a job while requiring humans to supervise, interpret and extend their outputs. For the individual employee, AI adoption therefore changes both the content of the job and the competencies required to perform it. This dual character underpins the mechanism proposed in this study.

2.2 Employee performance

Employee performance is conceptualised as the behaviours and outcomes, under the employee's control, that contribute to organisational goals. Following Motowidlo and Van Scotter (1994) and the Individual Work Performance Questionnaire (Koopmans et al., 2014), it comprises task performance, the proficiency with which core job tasks are completed in terms of quantity, quality and timeliness, and contextual performance, the discretionary behaviours that support the social and psychological environment of work, such as taking on challenging tasks, helping colleagues and proposing improvements. Both dimensions are plausibly affected by AI. Task performance may improve through faster information processing and fewer routine errors. Contextual performance may improve when AI frees time for collaboration and creative work, or it may suffer when AI isolates employees or erodes discretion (Budhwar et al., 2023).

2.3 Digital skills

Digital skills are the abilities that enable people to use digital technologies effectively for work and life. In a systematic review, van Laar et al. (2017) distinguished seven core digital skills (technical, information management, communication, collaboration, creativity, critical thinking and problem solving) and five contextual skills (ethical awareness, cultural awareness, flexibility, self-direction and lifelong learning). This multidimensional view is particularly relevant to AI. Operating an AI tool is a technical skill. Judging whether an AI-generated answer is accurate is an information-management and critical-thinking skill. Deciding how to combine AI outputs with one's own expertise is a problem-solving skill (Dell'Acqua et al., 2023; Raisch & Krakowski, 2021). From a human-capital perspective, digital skills are a form of general and firm-specific human capital that raises the marginal productivity of the employee (Becker, 1964), and their value increases with the technological intensity of the job.

A related concept is AI literacy, the ability to understand what AI systems can and cannot do, to use them effectively and to evaluate their outputs critically. AI literacy can be viewed as the AI-specific subset of the broader digital-skill repertoire. The information-management and critical-thinking components of digital skills are especially relevant to it, because generative AI systems produce fluent outputs whose correctness cannot be inferred from their form (Budhwar et al., 2023). The present study adopts the broader construct of digital skills for two reasons. First, AI tools are embedded in wider digital workflows, so their productive use depends on general digital competence. Second, validated multidimensional frameworks for digital skills are available, whereas measures of AI literacy at work

are still being developed. The DS items nevertheless include explicit references to AI outputs (DS2 and DS5) to capture the AI-specific evaluative component.

2.4 Theoretical framework

Four complementary perspectives inform the research model. The technology acceptance tradition (Davis, 1989; Venkatesh et al., 2003) explains why employees adopt AI but says little about the consequences of use. Task–technology fit theory (Goodhue & Thompson, 1995) closes this gap by proposing that technology improves individual performance when its functionality matches task requirements and when it is actually used. Importantly, the theory recognises individual characteristics, including skills, as determinants of fit. Human capital theory (Becker, 1964) explains why investments in employees' capabilities raise productivity and why technology and skills are complements. The job demands–resources (JD-R) model (Bakker & Demerouti, 2007) treats skills as personal resources that help employees cope with job demands, such as the cognitive demands created by new technology, and that foster engagement and performance.

Combining these perspectives yields a coherent account. AI adoption changes job demands and opportunities. Employees who use AI develop and exercise digital skills through repeated interaction, experimentation and feedback, which is a learning-by-doing process. These skills improve task–technology fit and act as personal resources, enabling employees to translate AI functionality into better task and contextual performance. AI may also affect performance directly, through automation of routine work, independently of skill development. The sociotechnical framework of Makarius et al. (2020) makes a similar argument at the organisational level, namely that AI creates value only when employees are 'socialised' into working with it, a process in which competence development is central.

2.5 Prior evidence on AI, skills and performance

Three strands of evidence are relevant to the proposed model. The first consists of controlled experiments and natural experiments on generative AI. These studies consistently find average productivity gains and also find that gains vary with worker characteristics. In customer support, AI assistance raised productivity by about 15 % on average, with much larger gains for novice and lower-skilled agents, partly because the AI system disseminated the tacit practices of high performers (Brynjolfsson et al., 2025). In professional writing, ChatGPT reduced completion time and raised quality while compressing the productivity distribution (Noy & Zhang, 2023). In consulting, the effect depended on whether the task lay inside or outside AI's capability frontier, and performance declined when consultants accepted incorrect AI output on tasks beyond that frontier (Dell'Acqua et al., 2023). These findings show that AI interacts with human capabilities, although the experiments typically measure short-term task performance and do not examine how capabilities change with sustained use.

The second strand is the organisational and HRM literature on AI implementation. Reviews identify the development of employee capabilities, alongside data, technology and governance resources, as a condition for realising value from AI (Chowdhury et al., 2023). Sociotechnical analyses argue that employees must be helped to understand and work with AI if it is to improve performance (Makarius et al., 2020). Conceptual work on generative AI in HRM highlights upskilling and reskilling as central HR responses (Budhwar et al., 2023; Tambe et al., 2019). This literature supports the mediating role of skills but has rarely tested it quantitatively. The third strand, labour economics, shows that technology changes the task content of jobs and raises the returns to complementary skills (Acemoglu & Restrepo, 2019; Autor, 2015). Taken together, the three strands point to the same mechanism, namely that AI's contribution to performance depends on and works through employees'

capabilities. They also show that an individual-level model linking adoption, skills and performance is still missing.

2.6 Hypotheses

Consistent with experimental and field evidence on AI-assisted work (Brynjolfsson et al., 2025; Noy & Zhang, 2023; Peng et al., 2023) and with the augmentation perspective (Raisch & Krakowski, 2021), AI adoption is expected to improve employee performance directly, for example by speeding up information retrieval, reducing routine errors and supporting decisions.

H1. AI adoption is positively related to employee performance.

Working with AI tools requires and develops technical, information-management and critical-evaluation skills. Employees who regularly use AI must learn to formulate requests, evaluate outputs and integrate them into workflows, and organisations that adopt AI typically accompany adoption with training (Chowdhury et al., 2023; Makarius et al., 2020).

H2. AI adoption is positively related to employees' digital skills.

Digital skills are personal resources and human capital that raise productivity in technology-intensive work (Bakker & Demerouti, 2007; Becker, 1964; van Laar et al., 2017). Employees with stronger digital skills can exploit digital and AI tools more fully and avoid their pitfalls (Dell'Acqua et al., 2023).

H3. Digital skills are positively related to employee performance.

Combining H2 and H3, digital skills are expected to transmit part of the effect of AI adoption to employee performance. Because AI may also improve performance through automation that does not depend on skill development, partial rather than full mediation is expected (Zhao et al., 2010).

H4. Digital skills mediate the positive relationship between AI adoption and employee performance.

Figure 1 presents the research model.

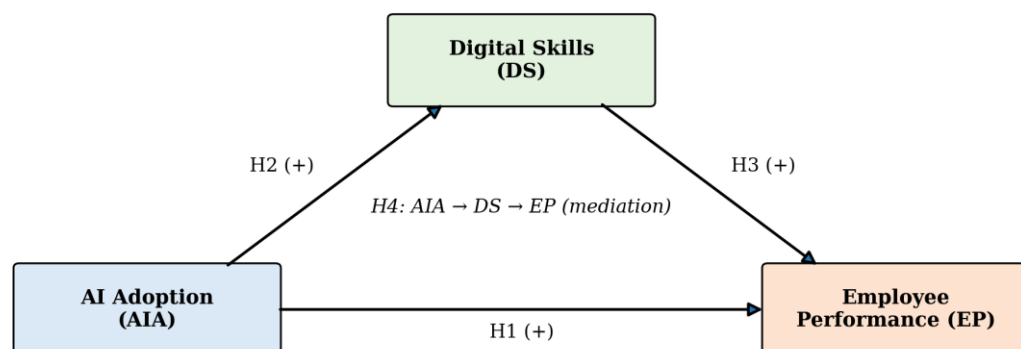


Figure 1. Research model and hypotheses.

3. Methodology

3.1 Research design

The study uses a Monte Carlo simulation design, structured according to the ADEMP framework (aims, data-generating mechanisms, estimands, methods and performance measures) recommended for simulation studies (Morris et al., 2019). Simulation is appropriate here for two reasons. First, it allows the full analytical workflow of a typical survey-based PLS-SEM study to be demonstrated and

documented transparently before field data are collected, which is a form of pre-registration of the analysis. Second, because the true population parameters are known, simulation is the only way to evaluate how accurately the analytical procedure recovers them, how much statistical power a given sample provides, and how sensitive the conclusions are to method bias. These questions cannot be answered with a single empirical dataset. The simulated data are not observations of real employees, and the estimates reported in Section 4 should be read as properties of the model and the method, not as empirical findings about any organisation.

3.2 Measurement instrument

Table 1 presents the 17-item instrument, measured on five-point Likert scales (1 = strongly disagree, 5 = strongly agree). AI adoption (AIA, five items) was adapted from the use-behaviour construct of UTAUT (Venkatesh et al., 2003) and from AI-capability research (Chowdhury et al., 2023). The items capture frequency of use, integration into workflows, availability and reliance on AI outputs. Digital skills (DS, six items) operationalise the core digital skills identified by van Laar et al. (2017): technical, information-management, communication and collaboration, creativity, critical-thinking and problem-solving skills, with an explicit reference to the evaluation of AI outputs. Employee performance (EP, six items) was adapted from the Individual Work Performance Questionnaire (Koopmans et al., 2014) and covers task performance (EP1–EP3) and contextual performance (EP4–EP6; Motowidlo & Van Scotter, 1994). All constructs are specified as reflective. In a field application, the instrument should be pre-tested and validated, and supervisor ratings or objective indicators of performance should be added where possible.

Table 1. Measurement instrument

Construct	Code	Item (5-point Likert scale)	Adapted from
AI adoption (AIA)	AIA1	I regularly use AI-based tools (e.g., generative assistants, predictive analytics) to perform my work tasks.	Venkatesh et al. (2003); Chowdhury et al. (2023)
	AIA2	AI applications are integrated into the core workflows of my job.	
	AIA3	My organisation provides me with access to AI applications relevant to my role.	
	AIA4	I use AI-generated outputs (recommendations, drafts, forecasts) when making work decisions.	
	AIA5	The extent to which I use AI in my work has increased over the past year.	
Digital skills (DS)	DS1	I can operate digital and AI tools confidently and solve common technical problems.	van Laar et al. (2017)
	DS2	I can search for, evaluate and verify digital information, including AI-generated content.	
	DS3	I can communicate and collaborate effectively with colleagues through digital platforms.	
	DS4	I can use digital tools to create new content or solutions.	
	DS5	I can recognise when the output of an AI tool is inaccurate or inappropriate.	

	DS6	I can quickly learn and adapt to new digital tools and applications.	
Employee performance (EP)	EP1	I complete my work tasks on time.	Koopmans et al. (2014); Motowidlo & Van Scotter (1994)
	EP2	The quality of my work meets high standards.	
	EP3	I achieve the results that are expected of me.	
	EP4	I take on challenging work tasks when they are available.	
	EP5	I help colleagues when they need support with their work.	
	EP6	I come up with creative solutions to work-related problems.	

3.3 Data-generating model

The population model (Table 2) specifies standardised latent variables with $AIA \rightarrow DS = a = .50$, $DS \rightarrow EP = b = .40$ and $AIA \rightarrow EP = c' = .15$. The implied population indirect effect is $a \times b = .20$ and the total effect .35. These values represent moderate effects of the size commonly hypothesised in technology-adoption and HRM research. They were fixed before any data were generated. Each item was generated as $\lambda \cdot \eta + e$, with standardised loadings between .70 and .84, and was then discretised into five ordered categories using thresholds that produce approximately symmetric response distributions with item-specific location shifts. The main sample comprised $n = 350$ cases, which exceeds the minimum suggested by the inverse-square-root method for a path of .15 at $\alpha = .05$ (Kock & Hadaya, 2018). A fixed random seed makes the dataset fully reproducible.

Table 2. Population (data-generating) model

Element	Specification
Structural paths	$AIA \rightarrow DS: a = .50$; $DS \rightarrow EP: b = .40$; $AIA \rightarrow EP: c' = .15$
Implied effects	Indirect $a \times b = .20$; total = .35; population $R^2(DS) = .25$, $R^2(EP) \approx .26$
Indicator loadings	AIA: .78, .82, .75, .80, .72; DS: .80, .77, .83, .74, .79, .71; EP: .81, .76, .84, .78, .73, .70
Response format	Five-point Likert categories from discretised continuous responses
Main sample	$n = 350$ (fixed seed); 5,000 bootstrap resamples
Experiment 1	1,000 replications, $n = 350$; PLS-SEM vs. PLSc
Experiment 2	$n = 50, 100, 150, 200, 300, 400, 500 \times 3$ effect scenarios $\times 400$ replications
Experiment 3	Method-factor loading $m = 0, .20, .30, .40 \times 400$ replications, $n = 350$

3.4 Analytical strategy

The data were analysed with PLS-SEM using the path-weighting scheme and Mode A estimation for reflective constructs (Hair et al., 2019). Because standard PLS-SEM estimates construct scores as weighted composites that contain measurement error, the model was also estimated with consistent PLS (PLSc). PLSc corrects the inter-construct correlations using the reliability coefficient ρ_A and thus estimates common-factor parameters consistently (Dijkstra & Henseler, 2015). The measurement model was evaluated through indicator loadings ($\geq .708$), internal consistency (Cronbach's α , ρ_A and composite reliability ρ_C , between .70 and .95), convergent validity (average variance extracted, $AVE \geq .50$) and discriminant validity (the Fornell–Larcker criterion and the heterotrait–monotrait ratio, $HTMT < .85$; Fornell & Larcker, 1981; Henseler et al., 2015).

The structural model was assessed through inner collinearity (VIF), path coefficients, coefficients of determination (R^2) and effect sizes (f^2 ; Cohen, 1988). Significance was evaluated with 5,000 bootstrap resamples and percentile confidence intervals (Efron & Tibshirani, 1993). Mediation was tested following Zhao et al. (2010), Preacher and Hayes (2008) and Hayes (2022), rather than with the causal-steps approach of Baron and Kenny (1986), which requires a significant total effect and has low power. The indirect effect $a \times b$ was considered significant if its 95 % bootstrap confidence interval excluded zero. The type of mediation (complementary, competitive or indirect-only) was then determined from the sign and significance of the direct effect. The variance accounted for ($VAF = \text{indirect}/\text{total effect}$) was reported as a descriptive measure of the strength of mediation. All procedures were implemented in Python (NumPy) from first principles. Construct scores, weights and loadings were checked against the standard PLS algorithm equations.

3.5 Monte Carlo experiments

Three experiments assessed the performance of the analytical strategy. Experiment 1 (parameter recovery) generated 1,000 samples of $n = 350$ from the population model and recorded the mean, standard deviation and root-mean-square error (RMSE) of the PLS and PLS_c estimates of a , b , c' and $a \times b$. Experiment 2 (statistical power) generated 400 samples for each combination of seven sample sizes ($n = 50$ to 500) and three effect-size scenarios: small ($a = b = .14$, $c' = 0$), medium ($a = b = .39$, $c' = 0$) and the hypothesised model. The values .14 and .39 correspond to the 'small' and 'medium' paths used by Fritz and MacKinnon (2007). Power was the proportion of samples in which the 95 % Monte Carlo confidence interval for the indirect effect excluded zero (Preacher & Selig, 2012). This procedure was used instead of the bootstrap to keep the 8,400 replications computationally tractable. Experiment 3 (common method variance) added a common method factor with loadings $m \in \{0, .20, .30, .40\}$ to every item (Podsakoff et al., 2003). For 400 samples per condition, it recorded the PLS estimates and the share of variance explained by the first unrotated factor, the quantity used in Harman's single-factor test.

4. Results

4.1 Measurement model

Table 3 reports the measurement model. All standardised loadings exceeded .69, and 16 of the 17 exceeded the .708 benchmark. EP6 (.694) was retained because the AVE of its construct remained above .50 and its content (creative problem solving) is theoretically important. Cronbach's α ranged from .852 to .875, ρ_A from .857 to .882 and composite reliability from .891 to .909, all within the recommended range of .70–.95. The AVE values (AIA = .667, DS = .608, EP = .576) supported convergent validity. Table 4 shows that the square root of each construct's AVE exceeded its correlations with the other constructs (Fornell–Larcker criterion), and all HTMT ratios were well below .85 (maximum .465 for AIA–DS). Discriminant validity was therefore established. In the simulated sample, Harman's single-factor test attributed 34.9 % of the item variance to the first factor.

Table 3. Measurement model: item statistics, loadings, reliability and convergent validity

Construct	Item	Mean	SD	Loading	α	ρ_A	ρ_C	AVE
AI adoption	AIA1	3.27	1.03	0.825	0.875	0.882	0.909	0.667
	AIA2	3.15	1.05	0.854				
	AIA3	3.05	1.06	0.789				
	AIA4	3.29	1.05	0.851				
	AIA5	3.20	1.03	0.763				
Digital skills	DS1	3.35	0.97	0.836	0.871	0.877	0.903	0.608

	DS2	3.25	0.98	0.749				
	DS3	3.12	1.03	0.808				
	DS4	3.34	1.00	0.764				
	DS5	3.22	0.99	0.775				
	DS6	3.19	0.99	0.742				
Employee performance	EP1	3.34	0.99	0.772	0.852	0.857	0.891	0.576
	EP2	3.17	0.98	0.728				
	EP3	3.11	1.02	0.831				
	EP4	3.29	1.02	0.770				
	EP5	3.17	1.06	0.753				
	EP6	3.14	1.01	0.694				

Note. Simulated data, $n = 350$. α = Cronbach's alpha; ρ_A = Dijkstra–Henseler reliability; ρ_C = composite reliability; AVE = average variance extracted.

Table 4. Discriminant validity: Fornell–Larcker criterion (lower triangle) and HTMT (upper triangle)

Construct	AIA	DS	EP
AI adoption (AIA)	0.817	0.465	0.297
Digital skills (DS)	0.412	0.780	0.427
Employee performance (EP)	0.261	0.374	0.759

Note. Bold diagonal values are the square roots of AVE; values below the diagonal are inter-construct correlations; values above the diagonal are HTMT ratios.

4.2 Structural model and hypothesis testing

Inner VIF values were 1.20, so collinearity among the predictors of EP was not a concern. Table 5 and Figure 2 present the structural results. AI adoption had a positive and significant effect on digital skills ($\beta = .412$, $t = 10.01$, $p < .001$, 95 % CI [.332, .495]), supporting H2, with a medium-to-large effect size ($f^2 = .205$). Digital skills had a positive and significant effect on employee performance ($\beta = .321$, $t = 6.43$, $p < .001$, CI [.227, .423]), supporting H3 ($f^2 = .101$, small to medium). The direct effect of AI adoption on employee performance was small but significant in standard PLS-SEM ($\beta = .129$, $t = 2.37$, $p = .025$, CI [.017, .235]; $f^2 = .016$), supporting H1. The model explained 17.0 % of the variance in digital skills and 15.4 % in employee performance. With PLSc the corresponding values were 22.0 % and 19.9 %.

Table 5. Structural model results and hypothesis tests (5,000 bootstrap resamples)

Hypothesis / path	β	SE	t	p	95 % CI	f^2	PLSc β	Decision
H1: AIA $\rightarrow$ EP (direct)	0.129	0.054	2.37	.025	[0.017, 0.235]	0.016	0.126 (n.s.)	Supported (PLS); not supported (PLSc)
H2: AIA $\rightarrow$ DS	0.412	0.041	10.01	< .001	[0.332, 0.495]	0.205	0.469	Supported
H3: DS $\rightarrow$ EP	0.321	0.050	6.43	< .001	[0.227, 0.423]	0.101	0.373	Supported
H4: AIA $\rightarrow$ DS $\rightarrow$ EP (indirect)	0.132	0.026	5.05	< .001	[0.088, 0.191]	–	0.175	Supported

Total effect AIA → EP	0.261	0.048	5.39	< .001	[0.168, 0.359]	—	—	—
--------------------------	-------	-------	------	--------	----------------	---	---	---

Note. $R^2(DS) = 0.170$; $R^2(EP) = 0.154$ (PLSc: 0.220 and 0.199). VAF = 50.7 %. Simulated data, $n = 350$.

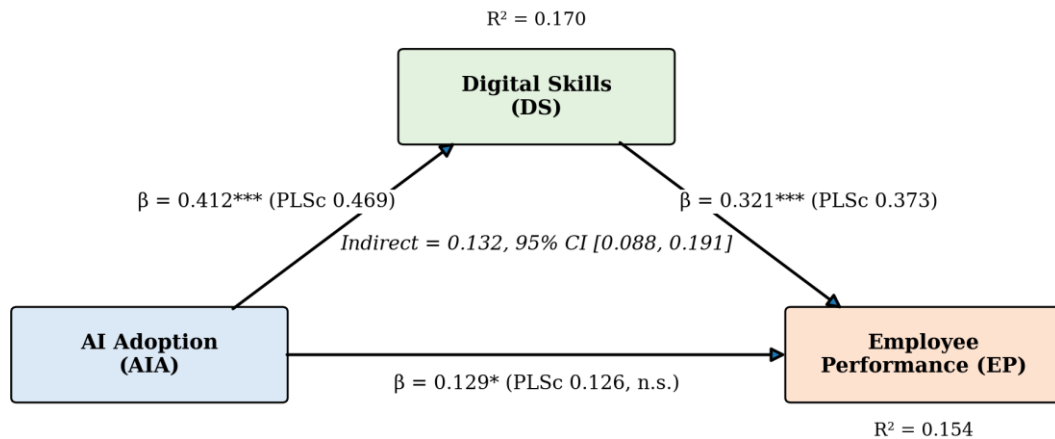


Figure 2. Estimated structural model (PLS-SEM; PLSc estimates in parentheses). * $p < .05$, * $p < .001$.**

The bootstrap distribution of the indirect effect (Figure 3) was approximately normal with a slight right skew. The indirect effect of AI adoption on employee performance through digital skills was $\beta = .132$ ($t = 5.05$, 95 % CI [.088, .191]). Because the interval excludes zero, H4 is supported. Together with the significant direct effect, this pattern indicates complementary partial mediation (Zhao et al., 2010). The total effect was .261 (CI [.168, .359]), and the indirect effect accounted for 50.7 % of it (VAF), which is at the upper boundary of partial mediation in conventional interpretations.

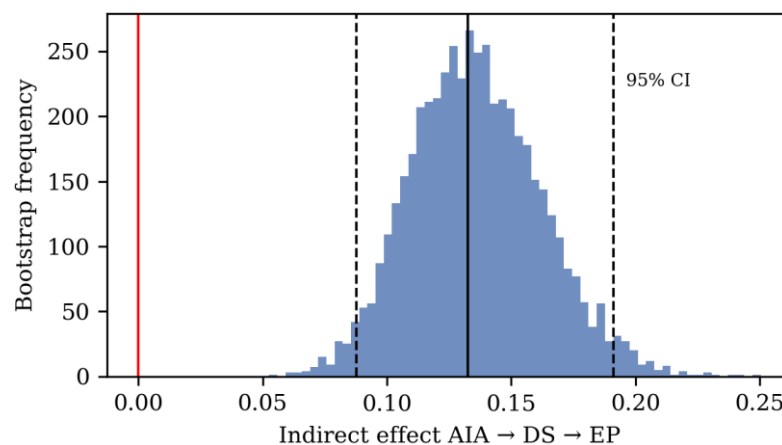


Figure 3. Bootstrap distribution of the indirect effect (5,000 resamples); solid line = point estimate, dashed lines = 95 % percentile confidence interval, red line = zero.

The PLSc estimates were systematically larger for the paths involving digital skills ($a = .469$, $b = .373$, indirect = .175, CI [.112, .255]). The direct effect, however, was no longer significant ($c' = .126$, $p = .071$, CI [−.013, .255]). Under the consistent estimator the evidence for H1 is therefore weaker, and the mediation would be classified as indirect-only. The two estimators thus lead to different substantive conclusions about the direct path. This divergence is examined systematically in Experiment 1.

4.3 Parameter recovery (Experiment 1)

Table 6 and Figure 4 summarise 1,000 replications. Standard PLS-SEM under-estimated the paths into and out of the mediator. The mean estimate of a was .436 against the true value of .50 (–12.8 %), and that of b was .353 against .40 (–11.8 %). The indirect effect was under-estimated by 23.1 % (.154 against .20). The direct effect c' was recovered almost exactly (.152 against .15), because attenuation of the a and b paths shifts explanatory weight back to the direct path. PLSc was virtually unbiased for all four parameters ($a = .500$, $b = .400$, $c' = .151$, $a \times b = .200$), at the cost of larger sampling variability (SD of $a \times b = .036$ against .026). In terms of RMSE, PLSc outperformed PLS-SEM for a (.047 against .077), b (.061 against .068) and the indirect effect (.036 against .053), whereas PLS-SEM had a slightly lower RMSE for c' (.054 against .067). The main sample was therefore typical: its PLS estimates of a (.412) and b (.321) were somewhat below the replication means, and its PLSc estimates fell within one standard deviation of the true values.

Table 6. Parameter recovery over 1,000 Monte Carlo replications (n = 350)

Parameter	True value	PLS mean (SD)	PLS bias (%)	PLS RMSE	PLSc mean (SD)	PLSc bias (%)	PLSc RMSE
a (AIA → DS)	0.50	0.436 (0.042)	-12.8	0.077	0.500 (0.047)	-0.0	0.047
b (DS → EP)	0.40	0.353 (0.049)	-11.8	0.068	0.400 (0.061)	-0.1	0.061
c' (AIA → EP)	0.15	0.152 (0.054)	+1.4	0.054	0.151 (0.067)	+0.6	0.067
$a \times b$ (indirect)	0.20	0.154 (0.026)	-23.1	0.053	0.200 (0.036)	-0.1	0.036

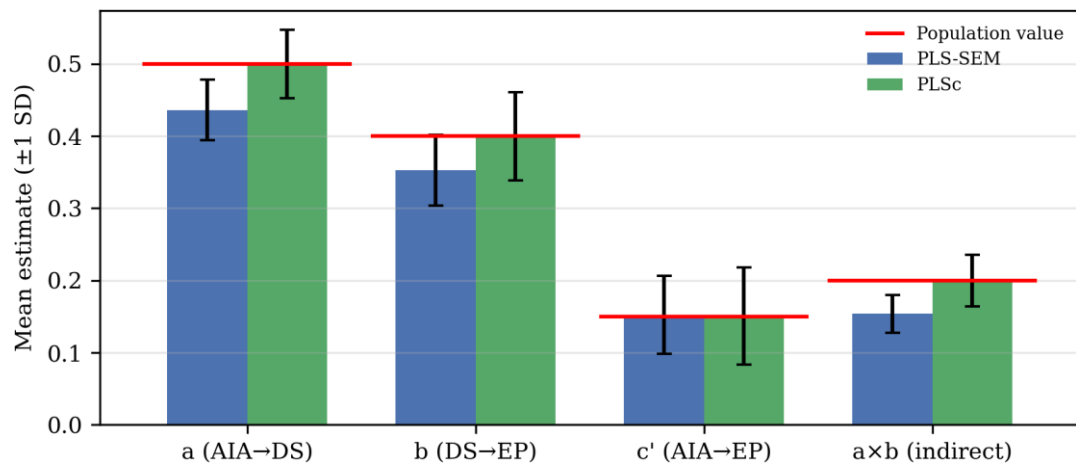


Figure 4. Mean PLS-SEM and PLSc estimates (± 1 SD) across 1,000 replications; red lines mark population values.

4.4 Statistical power (Experiment 2)

Table 7 and Figure 5 show power to detect the indirect effect. For the hypothesised model, power was .70 at $n = 50$, .94 at $n = 100$ and above .99 from $n = 150$. For medium paths ($a = b = .39$), power reached .91 at $n = 100$ and .98 at $n = 150$. For small paths ($a = b = .14$), power was only .30 at $n = 200$, .47 at $n = 300$, .56 at $n = 400$ and .76 at $n = 500$. More than 500 respondents would therefore be needed to reach the conventional 80 % level. These requirements are of the same order as, or higher than, the

sample sizes of about 500 that Fritz and MacKinnon (2007) reported for small paths in observed-variable mediation models. The attenuation of composite-based estimates documented in Experiment 1 reduces the effective effect size and therefore increases the sample needed.

Table 7. Power to detect the indirect effect (proportion of 400 replications with 95 % CI excluding zero)

Scenario	n = 50	n = 100	n = 150	n = 200	n = 300	n = 400	n = 500
Small ($a = b = 0.14$)	0.08	0.14	0.21	0.29	0.47	0.56	0.76
Medium ($a = b = 0.39$)	0.63	0.91	0.98	0.99	1.00	1.00	1.00
Hypothesised ($a = .50, b = .40, c' = .15$)	0.70	0.94	0.99	1.00	1.00	1.00	1.00

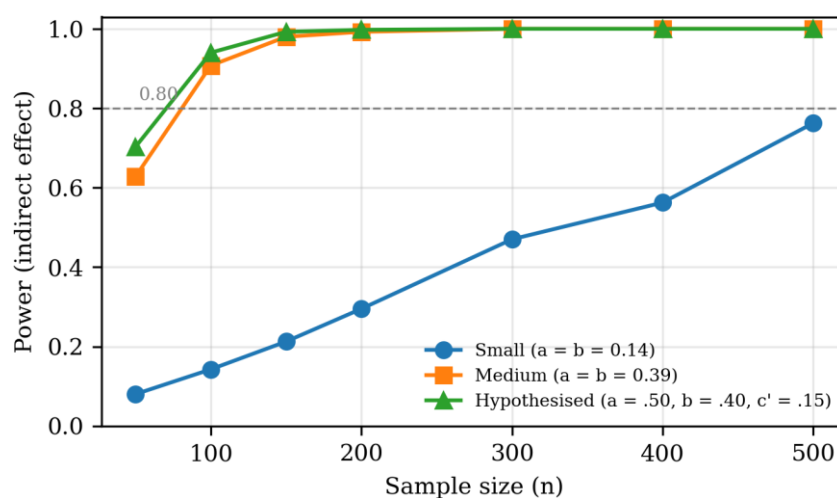


Figure 5. Statistical power for the indirect effect as a function of sample size.

4.5 Common method variance (Experiment 3)

Figure 6 and Table 8 show the consequences of common method variance. As the method-factor loading increased from 0 to .40, the mean PLS estimate of a rose from .435 to .559, b from .358 to .428, c' from .152 to .211 and the indirect effect from .156 to .239, an increase of 54 %. At $m = .30$, the indirect effect (.204) was inflated by 31 % relative to the method-free condition. It even happened to coincide with the true population value, because CMV inflation offset the attenuation from measurement error. The bias is therefore not detectable by inspecting the estimates. Harman's single-factor test was equally uninformative. The first factor accounted on average for 37.3 % of the variance without method bias, 40.7 % at $m = .20$ and 45.0 % at $m = .30$, all below the 50 % threshold commonly used to declare CMV absent. Only at $m = .40$ did the average share (51.0 %) exceed the threshold.

Table 8. Effect of common method variance on PLS estimates (400 replications per condition, $n = 350$)

Method loading	a	b	c'	$a \times b$	$\Delta a \times b$ (%)	First-factor variance (%)
0.00	0.435	0.358	0.152	0.156	+0.0	37.3 (max 43.1)
0.20	0.470	0.378	0.170	0.178	+14.2	40.7 (max 46.7)
0.30	0.510	0.401	0.190	0.204	+31.4	45.0 (max 50.3)
0.40	0.559	0.428	0.211	0.239	+53.7	51.0 (max 56.5)

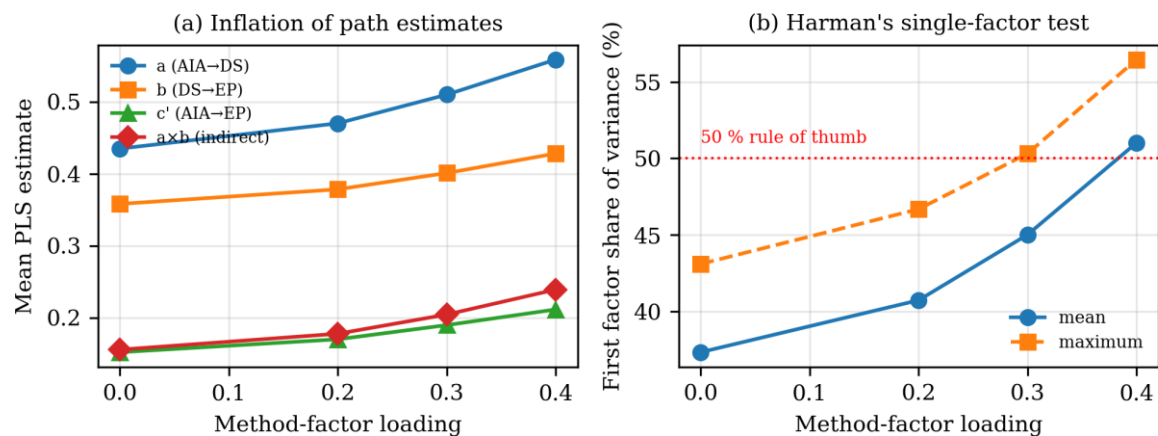


Figure 6. (a) Mean PLS estimates and (b) Harman's first-factor variance share as functions of the method-factor loading.

5. Discussion

5.1 Theoretical implications

This study proposed that digital skills are a key mechanism through which AI adoption affects employee performance. It specified the mechanism as a mediation model grounded in task–technology fit, human capital and JD-R theory. Under the specified model, roughly half of the total effect of adoption on performance travels through digital skills. The analysis shows how this mechanism would appear in the data a typical survey study collects and how it can be tested rigorously. Three theoretical implications follow.

First, the model shifts the focus of technology-acceptance research from adoption itself to the capabilities that adoption builds and requires. TAM and UTAUT explain why employees use technology (Davis, 1989; Venkatesh et al., 2003), but they are largely silent on why use translates into better work for some employees and not for others. Introducing digital skills as a mediator links the acceptance tradition to task–technology fit theory, in which individual abilities are part of the fit that determines performance impacts (Goodhue & Thompson, 1995). It thereby answers calls to study the outcomes of AI use rather than only its antecedents (Venkatesh, 2022).

Second, the model offers a micro-level explanation for the heterogeneous productivity effects of AI documented in recent experiments. Gains are largest where AI complements workers' capabilities and smallest, or negative, where workers lack the judgement to recognise the limits of AI output (Brynjolfsson et al., 2025; Dell'Acqua et al., 2023). The digital-skills pathway formalises this complementarity. It implies that the same level of adoption will produce different performance outcomes depending on how far adoption is accompanied by skill development. This is consistent with the automation–augmentation perspective of Raisch and Krakowski (2021) and with human capital theory's view that technology and skills are complements (Becker, 1964).

Third, the distinction between complementary partial mediation (under PLS-SEM) and indirect-only mediation (under PLSc) has theoretical significance. If the direct effect is genuinely small, as specified here, then most of the performance benefit of AI depends on employees' capabilities rather than on the technology alone. This is an 'augmentation-dominant' pattern. Conversely, a large direct effect would indicate an 'automation-dominant' pattern in which AI improves performance largely independently of employee skills. Empirical studies that report only standard PLS-SEM estimates risk

misclassifying the mechanism, because, as Experiment 1 showed, measurement error shifts explanatory weight from the indirect to the direct path.

The model is also consistent with the heterogeneity reported in the experimental literature, although it cannot be compared with it quantitatively because the simulated estimates are not empirical. In the experiments, workers with less pre-existing skill gained more from AI assistance on routine tasks (Brynjolfsson et al., 2025; Noy & Zhang, 2023), whereas performance on complex tasks depended on the ability to judge AI output (Dell'Acqua et al., 2023). These two findings fit the proposed model if the direct path captures the levelling effect of AI on routine work and the indirect path captures the skill-dependent gains on complex work. A field test could separate the two by measuring task complexity and modelling it as a moderator of both paths, a natural extension of the framework developed here.

5.2 Methodological implications

The Monte Carlo experiments yield four recommendations for empirical research on AI and work. (1) When constructs are measured reflectively with multiple items, consistent PLS should be reported alongside standard PLS-SEM, or covariance-based SEM should be used. Composite-based estimates under-stated the indirect effect by almost a quarter under conditions typical of well-validated scales (loadings .70–.84). (2) Sample-size planning should be based on the expected size of the indirect effect, not on heuristics such as ten times the number of arrows. A sample of about 150 is sufficient for moderate effects, but small paths require more than 500 respondents. Many published survey studies on AI adoption fall well below this threshold and are therefore underpowered for small mediated effects. (3) Harman's single-factor test should not be used as evidence that common method variance is absent. In Experiment 3, method variance large enough to inflate the indirect effect by 31 % left the first-factor share at 45 %, consistent with the criticisms of Podsakoff et al. (2003, 2012). Procedural remedies should be preferred: temporal separation of predictor and outcome measurement, multi-source data such as supervisor-rated or objective performance, and marker variables or a measured latent method factor. (4) Conclusions about full versus partial mediation should be stated with caution and checked against both estimators, because the classification can change with the estimator, as it did in the main sample.

5.3 Practical implications

For managers and HR professionals, the model implies that investment in AI technology should be matched by investment in people. If a substantial share of AI's performance benefit depends on digital skills, then deployments that provide tools without capability development will realise only part of their potential, a pattern consistent with organisations' reported difficulties in capturing value from AI (Chowdhury et al., 2023). Training should go beyond tool operation to address the skills that make AI use productive and safe: formulating tasks for AI systems, verifying and critically evaluating AI outputs, integrating them with domain knowledge, and recognising when AI should not be used (Dell'Acqua et al., 2023; van Laar et al., 2017).

Several HRM practices follow. Selection and promotion criteria can include digital and critical-evaluation skills. Learning-by-doing can be supported through pilot projects, communities of practice and protected time for experimentation. Performance management can recognise effective human–AI collaboration rather than only output volume. Because AI assistance appears to benefit less-experienced workers most (Brynjolfsson et al., 2025), skill-development programmes may also reduce performance gaps within teams. At the organisational level, the sociotechnical approach of Makarius et al. (2020) suggests that AI implementation projects should include explicit plans for employee socialisation and capability building. Policy-makers concerned with the labour-market effects of AI can likewise treat

digital-skill formation as a lever for turning automation into augmentation (Acemoglu & Restrepo, 2019; World Economic Forum, 2025).

5.4 Limitations and future research

The most important limitation is that the data are simulated. The population parameters were set a priori, and the study therefore cannot show that digital skills actually mediate the effect of AI adoption in any real organisation, or how large the effects are. What it shows is how the hypothesised mechanism would manifest in survey data, how well the standard analytical procedure recovers it, and how design choices affect the conclusions. The natural next step is a field study using the instrument in Table 1, ideally with a multi-wave design (adoption at time 1, skills at time 2, performance at time 3) and supervisor-rated or objective performance. Such a design would address both causal ordering and common method variance.

Second, the model is deliberately parsimonious. Plausible moderators include task routineness, which determines whether AI substitutes for or complements work; employees' trust in AI; organisational support and training; and individual characteristics such as age and experience. Additional mediators include self-efficacy, work engagement and job crafting. Possible adverse pathways, such as technostress, deskilling or over-reliance, may produce competitive mediation and should be modelled explicitly. Third, the simulation assumed multivariate-normal latent variables, equal-interval thresholds and a single method factor. Real survey data may exhibit non-normality, response styles and more complex method effects. Fourth, digital skills were treated as a single first-order construct. A higher-order model distinguishing technical, informational and critical-evaluation skills (van Laar et al., 2017) could reveal which skills matter most for AI-enabled performance. Finally, cross-cultural and sector comparisons would show whether the mechanism generalises across institutional contexts with different levels of digital infrastructure and training provision.

6. Conclusion

This study examined how AI adoption may improve employee performance and proposed that digital skills are a central mediating mechanism. Grounded in technology acceptance, task–technology fit, human capital and job demands–resources theory, the model was operationalised with a 17-item instrument and tested on simulated survey data of 350 employees using PLS-SEM, consistent PLS and bootstrapped mediation analysis. Under the specified model, the measurement properties were sound, and digital skills transmitted about half of the effect of AI adoption on performance (indirect effect .132, 95 % CI [.088, .191]). Monte Carlo experiments showed that standard PLS-SEM under-estimates the indirect effect by about 23 %, whereas consistent PLS recovers it without bias. They also showed that small mediated effects require more than 500 respondents, and that common method variance can inflate the indirect effect substantially without being detected by Harman's test. The study thus offers a theoretically integrated model, a ready-to-use measurement instrument and evidence-based methodological guidance for future empirical research. Its central message for practice is that AI creates value at work through people: the performance benefits of adoption depend on the digital skills employees bring to, and develop through, their work with AI. Confirming this mechanism with field data is the essential next step.

Declarations

Conflict of interest: The authors declare no conflict of interest.

Data availability: The simulated datasets and the Python code used for data generation and analysis are available from the corresponding author on reasonable request. All population parameters are reported in Table 2.

Ethics statement: No human participants were involved; the study uses simulated data only.

References

1. Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2), 3–30. <https://doi.org/10.1257/jep.33.2.3>
2. Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3), 3–30. <https://doi.org/10.1257/jep.29.3.3>
3. Bakker, A. B., & Demerouti, E. (2007). The Job Demands–Resources model: State of the art. *Journal of Managerial Psychology*, 22(3), 309–328. <https://doi.org/10.1108/02683940710733115>
4. Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
5. Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. Columbia University Press.
6. 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>
7. Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., Bamber, G. J., Beltran, J. R., Boselie, P., Cooke, F. L., Decker, S., DeNisi, A., Dey, P. K., Guest, D., Knoblich, A. J., Malik, A., Paauwe, J., Papagiannidis, S., Patel, C., Pereira, V., Ren, S., ... Varma, A. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. *Human Resource Management Journal*, 33(3), 606–659. <https://doi.org/10.1111/1748-8583.12524>
8. Chowdhury, S., Dey, P., Joel-Edgar, S., Bhattacharya, S., Rodriguez-Espindola, O., Abadie, A., & Truong, L. (2023). Unlocking the value of artificial intelligence in human resource management through AI capability framework. *Human Resource Management Review*, 33(1), Article 100899. <https://doi.org/10.1016/j.hrmr.2022.100899>
9. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
10. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
11. Dell'Acqua, F., McFowland, E., III, Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Kraye, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality (Working Paper No. 24-013). Harvard Business School.
12. Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly*, 39(2), 297–316. <https://doi.org/10.25300/MISQ/2015/39.2.02>
13. Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall.
14. Fornell, C., & Larcker, D. F. (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>
15. Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, 18(3), 233–239. <https://doi.org/10.1111/j.1467-9280.2007.01882.x>

16. Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, 19(2), 213–236. <https://doi.org/10.2307/249689>
17. Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (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>
18. Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3rd ed.). Guilford Press.
19. Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
20. Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal*, 28(1), 227–261. <https://doi.org/10.1111/isj.12131>
21. Koopmans, L., Bernaards, C. M., Hildebrandt, V. H., de Vet, H. C. W., & van der Beek, A. J. (2014). Construct validity of the individual work performance questionnaire. *Journal of Occupational and Environmental Medicine*, 56(3), 331–337. <https://doi.org/10.1097/JOM.0000000000000113>
22. Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research*, 120, 262–273. <https://doi.org/10.1016/j.jbusres.2020.07.045>
23. Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11), 2074–2102. <https://doi.org/10.1002/sim.8086>
24. Motowidlo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology*, 79(4), 475–480. <https://doi.org/10.1037/0021-9010.79.4.475>
25. 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>
26. Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of AI on developer productivity: Evidence from GitHub Copilot (arXiv:2302.06590). *arXiv*. <https://doi.org/10.48550/arXiv.2302.06590>
27. Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63, 539–569. <https://doi.org/10.1146/annurev-psych-120710-100452>
28. Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
29. Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891. <https://doi.org/10.3758/BRM.40.3.879>
30. Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures*, 6(2), 77–98. <https://doi.org/10.1080/19312458.2012.679848>

31. Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192–210. <https://doi.org/10.5465/amr.2018.0072>
32. Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review*, 61(4), 15–42. <https://doi.org/10.1177/0008125619867910>
33. van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577–588. <https://doi.org/10.1016/j.chb.2017.03.010>
34. Venkatesh, V. (2022). Adoption and use of AI tools: A research agenda grounded in UTAUT. *Annals of Operations Research*, 308(1–2), 641–652. <https://doi.org/10.1007/s10479-020-03918-9>
35. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>
36. World Economic Forum. (2025). The future of jobs report 2025. World Economic Forum.
37. Zhao, X., Lynch, J. G., Jr., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37(2), 197–206. <https://doi.org/10.1086/651257>.