EUROPEAN INTERNATIONAL UNIVERSITY

EIU Journal of Action Research

Statistical Evaluation of AI-Enabled Training Micro-Agents: A Longitudinal Analysis of Adoption and Learning Efficiency

By Dr Smrite Goudhaman

Abstract

This study undertakes a longitudinal evaluation of the effectiveness of AI-enabled training within a frontline hospitality context, which is an extension of the preceding pilot study designed within a doctoral research context. The pilot study, designed as a controlled study, recruited 100 frontline employees between September and November 2024, evaluating learning outcomes within a supervised context. Following the completion of the doctoral research, the AI-enabled training intervention was rolled out within the operational context during 2025, allowing for a longitudinal evaluation of learning outcomes within a real-world context.

The independent variable is the mode and maturation of AI-enabled training, which is operationalized as a transition between a supervised pilot study and a rollout of AI-enabled training within an operational context, utilizing AI micro-agents and human supervision. The dependent variables are learning adoption and learning efficiency, which are operationalized as objective learning-platform trace data, including: course completion, assessment performance, and time-on-task. The study context is characterized by high workforce turnover, making learning efficiency an important dependent variable, where an exposure-adjusted active employee framework is employed to minimize bias and identify patterns among active employees.

Descriptive statistics show that the completion percentage was 100% for the pilot phase, where all the active employees (656 records) completed the process. However, the percentage remained stable at 86.82% (or 8,505 records out of 9,796 records) for the 2025 scaled deployment period. The z-test revealed a significant difference between the two percentages, where the z-score was 11.52, and the difference was 13.18% ($p < 0.001$), corresponding to a 95% CI of [12.51, 13.85].

However, the most significant effect of the scaled deployment was the improvement in learning efficiency, where the mean assessment score increased from 80.10 (SD = 21.55) for the pilot phase to 83.38 (SD = 23.14) for the 2025 deployment, $t(1108.70) = 4.29$, $p < 0.001$, and a small effect size of 0.14, 95% CI = [1.78, 4.79]. The time-on-task also reduced from 10.30 minutes (SD = 11.53) for the pilot phase to 5.98 minutes (SD = 7.82) for the 2025 deployment, $t(971.64) = -10.88$, $p < 0.001$, and a medium effect size of -0.52, 95% CI = [-5.09, -3.53].

This study contributes to the rare research on the performance of AI micro-agents in the context of the doctoral pilot and its longitudinal effects on the post-dissertation period, providing a measurement framework for the evaluation of the effectiveness of AI-enabled learning systems.

Keywords: AI agents • Human-in-the-loop governance • Learning analytics • Longitudinal study • Training effectiveness • Technology acceptance • Trust calibration • Frontline workforce • Hospitality operations • Responsible AI

1. Introduction

In the realm of service industry organizations, digital training platforms have become an important tool for training front-line workers. These environments are characterized by time pressure, role variability, and constant workforce churn. These factors make it difficult for workers to develop and sustain their skills. Despite the initial enthusiasm and positive engagement with training programs, most organizations have faced challenges in determining whether learning outcomes are sustained as training programs move from controlled pilots into the natural workflow. In recent times, AI-based training platforms have been proposed as a solution for addressing the challenges faced by front-line workers. These platforms have been designed to utilize the benefits of artificial intelligence by incorporating it into learning workflows. This has been proposed as a solution for improving learning outcomes and engagement. Despite the positive potential of AI-based training platforms, most studies have been conducted on their initial performance. This has left a knowledge gap regarding whether learning outcomes are sustained as training programs move from pilots into the natural workflow. Past research on digital training and information systems has focused on the initial performance and post-adoption behavior. Despite the positive potential of AI-based training platforms, most studies have been conducted on their initial performance. This has left a knowledge gap regarding whether learning outcomes are sustained as training programs move from pilots into the natural workflow. Past research on digital training and information systems has focused on the initial performance and post-adoption behavior. Despite the positive potential of AI-based training platforms, most studies have been conducted on their initial performance. This has left a knowledge gap regarding whether learning outcomes are sustained as training programs move from pilots into the natural workflow. Past research on digital training and information systems has focused on the initial performance and post-adoption behavior.

This lack of longitudinal data also brings theoretical and practical concerns. From a theoretical standpoint, the lack of longitudinal data brings concerns regarding the overall learning systems and how they might adapt and change as the organization and its conditions continue to fluctuate.

1.1 The Effectiveness of Training beyond the Pilot

Classic studies on the effectiveness of training emphasize the importance of learning and how the learning process impacts the overall organization, including the need for reinforcement and the overall ability to apply what was learned (Baldwin & Ford, 1988). The pilot, while a critical component, also brings concerns regarding the overall environment and how it might compare to the overall organization.

This can become a significant concern, especially within a service environment, where the employee might not only drop out of the training process but also might not begin the training process at all, or might not make the learning process a priority due to the overall environment and the need to perform the job.

1.2 The Effectiveness of Technology beyond the Pilot

Classic studies on the overall effectiveness of technology and its adoption by the organization and its employees also emphasize the importance of the learning process and the overall environment, including the need to continue using the information system (Davis, 1989; Venkatesh & Davis, 2000; Venkatesh et al., 2003).

Although such models offer valuable insights into the initial period of adoption, they offer limited guidance on the evolution of patterns and their implications for long periods. In a training context, initial compliance does not always ensure long-term engagement or learning outcomes. Thus, a longitudinal analysis is important to separate the effects of adoption and learning outcomes.

1.3 AI-enabled training, micro-agents, and human oversight

Advances in AI-enabled human resource and learning management have led to the development of micro-level interventions that aim to guide user behavior. These micro-level AI-enabled micro-agents aim to support users through reminders and prompts that help users complete tasks and learn. These micro-level interventions aim to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions aim to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions aim to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions aim to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions aim to support learning and performance by reducing cognitive and operational barriers.

Although such micro-level interventions have been designed to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions have been designed to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions have been designed to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions have been designed to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions have been designed to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions have been designed to support learning and performance by reducing cognitive and operational barriers.

Although such micro-level interventions have been designed to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions have been designed to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions have been designed to support learning and performance by reducing cognitive and operational barriers. These micro-level interventions have been designed to support learning and performance by reducing cognitive and operational barriers.

In training settings, these issues are particularly pertinent. Learning outcomes have significant implications for service quality, safety, and efficacy. Therefore, AI-based training systems must be assessed from a perspective that extends beyond efficiency to governance, i.e., how AI-based micro-agents interact with human-in-the-loop systems that ensure training integrity.

1.4 From Doctoral Pilot to Longitudinal Evidence

The present research extends from previous doctoral research into AI-based training in frontline hospitality settings. A controlled pilot study was conducted among 100 frontline workers between September and November 2024 to measure initial adoption, learning, and feasibility under controlled conditions.

Subsequent to completing the dissertation, the organization continued the same training program as part of their regular training processes in 2025. This provided a singular opportunity to examine how learning outcomes are affected after training extends beyond the confines of the doctoral pilot study into a training environment beyond initial conditions. In other words, the pilot study was not considered in isolation but as a precursor to longitudinal analysis of training system maturity beyond the pilot.

1.5 Rethinking Learning Outcomes in High-Churn Environments

Training effectiveness in frontline environments poses specific challenges due to high employee churn rates. Employees may leave the organization prior to completing their allocated training, and new employees may join the organization midway through the training process. Learning systems record all training interactions, including those that are incomplete, which can skew overall results if exposure is not factored into the analysis.

To address this problem, this current research proposes an alternative by using an "exposure-adjusted" approach, where only active employees, i.e., those with evidence of learning activity over given periods of time, are considered. This allows for a contextual interpretation of completion, with a focus on the quality and efficiency of learning, i.e., assessment performance and time on task. These objective measures provide a stronger basis for assessing learning outcomes than subjective self-report measures, which are subject to a variety of method biases (Podsakoff et al., 2003).

1.6 Study purpose, hypotheses, and contributions

This study aims to explore the impact of AI-enabled training outcomes as the program moves from a pilot phase to a scaled, longitudinal deployment phase. More specifically, the study aims to explore the impact of learning adoption and learning efficiency as the program moves from the pilot phase to the scaled deployment phase, where AI-enabled micro-agents are used to operate under a human-in-the-loop governance framework.

Informed by the theoretical discussion presented above, the following research hypotheses are proposed:

H1: Course completion rates will be different between the supervised pilot phase and the scaled deployment phase.

H2: Mean assessment performance will be different between the supervised pilot phase and the scaled deployment phase.

H3: Mean time-on-task will be different between the supervised pilot phase and the scaled deployment phase.

Informed by the research questions, the study operationalizes deployment maturity as the independent variable, learning adoption, and learning efficiency as the dependent variables, allowing the study to explore the post-pilot performance of the system.

This paper makes three contributions to the digital information systems field. First, the study presents longitudinal, post-dissertation research on AI-enabled training effectiveness in a real-world, frontline context. Second, the study illustrates the significance of churn-aware evaluation approaches to interpreting learning adoption and learning efficiency outcomes. Third, the study extends the understanding of AI micro-agents, human collaboration, and learning performance beyond the pilot phase.

1.7 Structure of the paper

This paper is structured as follows. Section 2 reviews the literature on training effectiveness, technology adoption, AI-enabled learning, and human-AI collaboration. Section 3 presents the research model and research hypotheses. Section 4 presents the research design, while Section 5 presents the research findings. Section 6 presents the discussion, while Section 7 concludes the paper.

2. Literature Review

2.1 Training Effectiveness and the Transfer Problem Over Time

Previous literature on training effectiveness has long recognized that the effectiveness of learning outcomes cannot be evaluated in a single point of time. Baldwin & Ford's foundational research on transfer of training has recognized that the effectiveness of learning outcomes over time is dependent on how learning is reinforced, applied, and supported over time. Though the effectiveness of learning outcomes may be high in the short term, effectiveness tends to diminish over time as organizational conditions change.

This is particularly the case in frontline settings, where employees are under time pressure, exhibit role variability, and have high turnover. In such settings, it is critical that training systems not only provide learning outcomes but also sustain employee engagement and learning behavior as employees join and leave the organization. Hence, longitudinal evaluation is critical in evaluating the effectiveness of learning outcomes, rather than single-point evaluation.

However, despite considerable research on the transfer of learning outcomes, few studies have longitudinally evaluated learning outcomes in pilot and field studies. Most studies have evaluated the effectiveness of learning outcomes either shortly after the completion of the learning program or within a very narrow time window.

2.2 Technology Acceptance and Adoption Beyond Initial Use

The effectiveness of digital learning systems has also been extensively examined through technology acceptance models. Davis' foundational research on technology acceptance has recognized that "perceived usefulness" and "perceived ease of use" are critical factors in determining the effectiveness of technology adoption. Subsequent extensions, such as TAM2 and UTAUT, have recognized the role of social factors, experience, and facilitating conditions in determining the effectiveness of technology adoption.

Although such models offer valuable insights into the initial period of adoption, they are less explicit about the dynamics of usage patterns as the system evolves. For instance, an individual may comply with a training program during the pilot period due to its visibility and supervision. However, their patterns of adoption may change as the level of supervision decreases. This differentiation is particularly important when considering training programs where initial compliance with a program may be a result of structural supervision rather than intrinsic engagement. Thus, a longitudinal approach must be taken to understand the differentiation between initial compliance and eventual patterns of adoption. Without such differentiation, an organization may overestimate the efficacy of a training program based on initial patterns of engagement.

2.3 AI-Enabled Training and Micro-Level Interventions

Significant advancements have been made in AI-Enabled Human Resource and Learning Systems. These advancements have led to the development of micro-level interventions that can support user behavior. For instance, micro-level interventions can be used as reminders and cues that can be directly incorporated into workflows. These micro-level interventions are designed to support user behavior. They are not intended to replace decision-making but rather reduce the level of friction and encourage users to engage with a desired course of action.

Typology of AI-Enabled Micro-Agents

To understand the conceptual scope of micro-level AI interventions within a training context, micro-agents can be classified into four main types:

1. Nudging Agents: Agents that encourage user engagement with a workflow or a specific activity.
2. Sequencing Agents: Agents that can be used as a sequencing tool that structures learning and can be used to determine the order and time taken.
3. Reinforcement Agents: Agents that can be used as a reinforcement tool that can be used to reinforce learning.
4. Friction-Reduction Agents: System-level changes that reduce cognitive or operational hurdles (e.g., auto-saving, micro-quizzes, easy navigation).

These categories are related to behavioral reinforcement theory and digital workflow design best practices, providing insight into the way AI micro-agents function at the task level, rather than at the strategic level. Thinking about micro-agents from this structured perspective allows for more precise analysis when examining the long-term effects of micro-agents on employees.

In the learning context, these micro-agents have the potential to influence the way employees utilize training materials, potentially increasing the likelihood of successful completion of training. However, the effectiveness of these micro-agents is not just dependent on the algorithms used to program the micro-agents, but the way employees perceive the AI micro-agents.

Algorithm aversion research suggests that employees are more likely to be dissatisfied with, and even terminate, interactions with algorithms after witnessing errors or inconsistencies (Dietvorst et al., 2015).

2.4 Human-AI collaboration and the role of oversight

An emerging body of research supports the position that AI micro-agents are best implemented as tools that complement human judgment, rather than as tools that function independently. Research on human-AI collaboration highlights the cognitive challenges that occur when humans and AI micro-agents are not properly defined

(Fügener et al., 2019). More recent research supports the position that AI micro-agents are best implemented as tools that complement human judgment, where the strengths of the human are augmented by the AI micro-agents, while the human is responsible for providing the necessary oversight (Hemmer et al., 2025).

Guidelines for human-AI interaction emphasize the importance of keeping the human informed, allowing the human to intervene, and providing mechanisms for accountability (Amershi et al., 2019).

In the context of AI-enabled training, human oversight plays several roles, including monitoring learner progress, handling exceptions, reinforcing expectations, and ensuring that AI-enabled nudges are consistent with organizational objectives. In the absence of human oversight, the interventions of micro-agents are likely to be ineffectual or even counterproductive.

2.5 Integrative Synthesis: Training, TAM, and Human-AI Systems

The literature on training effectiveness, technology adoption, and human-AI collaboration has largely developed in parallel, but their integration is necessary to achieve a full understanding of AI-enabled training systems.

Training effectiveness literature emphasizes reinforcement and transfer mechanisms (Baldwin & Ford, 1988). Technology acceptance literature explains the underlying determinants of initial adoption and long-term usage patterns (Davis, 1989; Venkatesh & Davis, 2000; Venkatesh, Morris, & Davis, 2003). Finally, human-AI collaboration literature emphasizes the importance of governance and oversight structures necessary to support effective human-AI collaboration (Fügener et al., 2019; Hemmer et al., 2025; Amershi et al., 2019).

AI-enabled training systems are located at the intersection of all three fields of study. Micro-agents are the vehicle for reinforcement mechanisms at scale. Technology acceptance factors influence engagement with reinforcement interventions. Finally, human oversight determines whether algorithmic nudges are trusted, calibrated, and consistent with organizational objectives.

Longitudinal AI-enabled training effectiveness relies on three interrelated mechanisms: behavioral reinforcement (training transfer theory), technology adoption and usage (TAM/UTAUT), and governance and oversight complementarity (human-AI systems theory).

2.6 Measuring Learning Outcomes: Beyond Completion Rates

Much of the training literature relies on self-reported data or single factors like completion rates. Nevertheless, these approaches are susceptible to common method bias and may not adequately capture learning outcomes (Podsakoff et al., 2003). Tractable learning platform data, like assessment scores and time on task, provide more reliable data for assessing learning effectiveness.

Assessment scores give insight into learning effectiveness, and time on task provides information about learning efficiency. A reduction in time on task with stable or improved assessment scores indicates increased learning efficiency and not lack of engagement (Pappas et al., 2019). Such data is especially relevant in longitudinal studies where changes in learning efficiency may indicate maturation of learning systems and users' familiarity.

In the frontline environment characterized by high workforce churn, it is equally important to account for exposure effects. Assessing learning effectiveness for employees who have remained engaged avoids bias from partial engagement and ensures more accurate interpretation of results.

2.7 Positioning the present study

The present study contributes to these streams of literature by addressing three gaps. First, it extends the literature on training and technology adoption studies by providing longitudinal data that extends beyond pilot studies and post-supervision system maturation. Second, it extends the literature on AI-enabled learning studies by empirically evaluating categorized micro-agent interventions in a human-in-the-loop governance framework. Third, it extends methodological literature by using a churn-aware and exposure-adjusted approach in a frontline environment characterized by high workforce turnover.

The combination of longitudinal data from a doctoral pilot and post-dissertation longitudinal data provides a unique opportunity to investigate how AI-enabled training systems mature, normalize, and sustain learning performance in real-world conditions.

3. Research Model and Research Questions

3.1 Conceptual framing

The goal of the current research is to examine the evolution of learning outcomes as an AI-supported training system progresses from a doctoral pilot to a larger operational deployment. Rather than examining the effectiveness of various training interventions, the current research follows the same AI-supported training system over time, enabling the evaluation of the system's learning outcomes as the deployment conditions mature.

In this context, the current research adopts a within-system longitudinal research design in which the primary explanatory factor is not the type of technology but the maturity of the system's deployment. This conceptual framing is consistent with previous research in the information system and training literatures, in which the primary concern has been the transfer of learning, long-term use, and performance outcomes over time rather than point-in-time adoption (Baldwin & Ford, 1988; Venkatesh et al., 2003).

The research model centers on the relationship between the maturation of AI-supported training, facilitated by micro-agents and governed via human oversight, and the observable learning outcomes for frontline employees.

3.2 Independent variable: Operationalizing deployment maturity

In the current research, the independent variable of interest is the maturity of the AI-supported training system's operational deployment.

To move beyond a narrative description of the research model and ensure clarity of the operationalization of the independent variable, the current research operationalizes the independent variable as a binary phase indicator variable as follows:
Deployment Maturity = 0 for the Doctoral Pilot Phase (September – November 2024)
Deployment Maturity = 1 for the Scaled Operational Phase (January – December 2025)

This binary operationalization of the independent variable reflects a change from a structured and supervised setting (pilot) to a naturalistic and widespread setting (scaled).

This process is in line with earlier studies on post-adoption system maturation and usage transitions (Venkatesh & Davis, 2000). It allows for a statistical evaluation of learning outcomes under deployment conditions, with the underlying configuration of the AI system being constant.

3.3 Dependent Variables

Three dependent variables are used in this study, all of which are based on objective data from the learning platform's usage traces. These cover both aspects of learning adoption and learning efficiency, addressing criticisms of evaluation methods focusing on a single metric.

3.3.1 Learning Adoption - Completion Rate

Learning adoption is measured by completion rate, defined as:

Completion Rate = Completed Records / Assigned Records

This is a direct measurement of the proportion of assigned course records that employees complete.

Completion is interpreted in a contextual manner, with pilot completion being a measure of structured supervision and scaled completion being a measure of actual workforce engagement in a real-world environment.

To minimize potential distortions from workforce churn, completion is measured among active employees, defined by sustained learning activity within a specified phase window. This ensures that completion is a fair reflection of actual participation, rather than being confounded by attrition effects.

3.3.2 Learning Quality - Assessment Performance

Learning quality is measured by mean assessment scores, defined as:

Mean Assessment Score = Sum of Assessment Scores / Completed Records

Assessment scores provide a direct measurement of actual learning outcomes, avoiding self-reported biases and mitigating common method biases (Podsakoff et al., 2003).

Changes in assessment performance between deployment phases are interpreted as changes in learning quality with deployment maturity.

3.3.3 Learning Efficiency - Time-on-Task

Learning efficiency is measured by mean time-on-task, defined as:

Mean Time-on-Task = Sum of Time-on-Task / Completed Records

This is a direct measurement of the average time in minutes spent by employees completing course material.

Time-on-task is viewed together with assessment performance. If time-on-task decreases while assessment performance is stable or increasing, this suggests that the child is learning more quickly rather than being unengaged (Pappas et al., 2019). Together, assessment performance and time-on-task provide a balanced and behaviorally grounded view of learning effectiveness.

3.4 Role of AI micro-agents and human oversight

In both training phases, the training system was configured to include AI-enabled micro-agents that support learning effectiveness. However, the configuration of the micro-agents was substantively similar across both training phases. What differed was the organizational context of their operation.

The system was configured to operate within a human-in-the-loop governance model in which managers and training leaders are accountable for monitoring progress, managing exceptions, and ensuring accountability.

The configuration of the system follows established guidelines on effective human-AI interaction that emphasize transparency, human control, and accountability (Amershi et al., 2019) as well as the complementarity of AI systems and human judgment (Fügener et al., 2019; Hemmer et al., 2025).

Clarification of the human oversight construct. Human oversight was conceptualized as a contextual governance condition rather than a moderator variable that was empirically determined in the model. Oversight levels were structurally higher in the pilot compared to the scaled deployment, but a quantitative measure of the intensity of human oversight was not available. Oversight was therefore conceptualized as a contextual feature of training maturity levels. The clarification of the human oversight construct ensures that it is conceptually precise, removing any potential issues of validity in relation to the measurement of the construct.

3.5 Hypotheses

Based on the operationalized model, the study tests the following hypotheses:
H1: Course completion rates are different between the supervised pilot phase (Deployment Maturity = 0) and the scaled operational phase (Deployment Maturity = 1).
H2: The mean assessment performance varies between the supervised pilot phase and the scaled operational phase.
H3: The mean time-on-task varies between the supervised pilot phase and the scaled operational phase.

These research hypotheses allow for the empirical verification of the effect of learning adoption and efficiency, as the AI-enabled training process shifts from the supervised pilot phase to the scaled operational phase.

3.6 Summary of the research model

The conceptual research model, as represented in Figure 1, was used to guide the development of the research design and the subsequent empirical verification.

The independent variable, as represented by the research model, is the deployment maturity of the AI-enabled training process, represented by a binary indicator, where the supervised doctoral pilot phase equals 0 and the scaled operational phase equals 1.

The deployment maturity of the AI-enabled training process was used to explain the variations in the three dependent variables, represented by the learning platform's objective data, including learning adoption, learning quality, and learning efficiency.

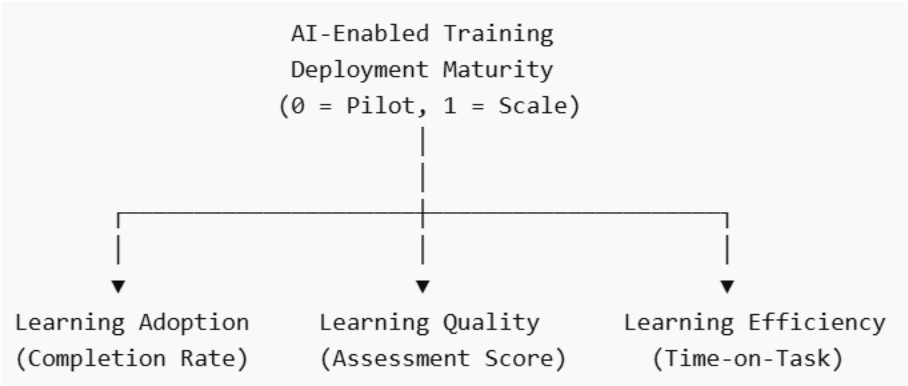

Figure 1 presents the conceptual research model.

Figure 1: Research model depicting the association between the AI-enabled training deployment maturity and the learning process. The micro-agents, designed to support the behavior of the learners, are the same in both phases, using nudging, sequencing, and reinforcement mechanisms. However, human oversight, as a moderating effect, was not considered a significant factor, as it was embedded within the structure of the two phases. The research model, therefore, tested the effect of the deployment maturity conditions on the learning process, while the configuration of the AI system was held constant. This structure aligns the conceptual approach, the variables, and the statistical verification, as represented in the inferential statistical tests, as discussed in Section 5.

4. Research Design and Methodology

4.1 Research Design

This study utilizes a longitudinal cohort extension design to examine the impact of AI-enabled training micro-agents on AI adoption, learning efficiency, and human management. It is a sequel to the pilot study conducted in 2024, where the design is extended to include a scaled deployment in 2025 to allow for the examination of the impact across two temporally distinct periods, while maintaining consistency in training objectives, platform architectures, and management controls.

In this study, AI-enabled training is conceptualized not as a simple instructional intervention, but rather as a set of constrained task-specific micro-agents that are embedded into the organizational workflow, monitoring learner response patterns while implementing pre-specified interventions, while at the same time working within boundaries that are explicitly defined, supervised, and controlled by human managers. The longitudinal design is particularly pertinent to this study, given the expected evolution of the behavioral impact of the micro-agents as the intervention becomes more embedded into the organizational workflow.

It is noteworthy that the study utilizes a non-experimental research design, where no experimental manipulation is implemented. Therefore, the study findings are best interpreted as longitudinal associations rather than causal effects.

4.2 Organizational Context: Toscano

Toscano is the organizational context of this study, where the study is set in the context of the casual dining restaurant business, operating in a high-pressure service context. In the context of the pilot study conducted in 2024, Toscano is planning to expand from the existing 35 outlets to 50 outlets, creating high operational pressures to scale up the training of the frontlines while maintaining consistency in service quality, food safety, and customer experiences.

This expansion context is significant from the perspective of analysis, as rapid outlet expansion usually accentuates the training issues that are common within the hospitality industry, including the quality of onboarding, skills gaps between staff and specific outlets, and the number of first-time managers. When Toscano was set to expand its outlets, the organization recognized the importance of training its front-line staff, specifically the need for training consistency, which was considered a critical risk factor that cannot be effectively managed using manual methods.

4.3 AI-Enabled Training Platform: SafetyCulture

To address the training issues associated with rapid expansion, Toscano decided to pilot the use of SafetyCulture, an AI-enabled digital training and operations platform, in 2024. The pilot group consisted of 100 front-line staff from some of the organization's outlets.

As a digital training platform, SafetyCulture was used by Toscano as an AI-enabled training environment, where the organization was able to use structured learning modules, automated task assignment, and training analytics to track the performance and behavior of the front-line staff. While the platform was able to generate recommendations, there was no decision-making authority exercised by the platform, and all decisions regarding the interpretation of the behavior and performance of the staff were left to the human managers.

Following the completion of the pilot phase, Toscano expanded its scale in 2025 in association with its outlet expansion, spreading the AI-based training system further through its employee population. Notably, the content of the training and micro-agents were kept unchanged in order to facilitate longitudinal comparison. The scale phase represented a shift from experimental use to organizational use of the system, thereby creating an environment where its use would naturally be viewed in terms of its evolution from new to familiar.

4.4 Data Sources and Cohort Definition

The primary source of data was platform-generated learning trace data, obtained from SafetyCulture from September 2024 to December 2025.

Cohort Clarification

The initial number of employees enrolled was around 100.

Employees who were considered active were those who had shown some level of engagement, i.e., at least one assignment record was completed.

Using the above exposure-adjusted cohort definition, the following was observed:

- 65 employees were considered active during the 2024 pilot window.
- 250 employees were considered active during the 2025 scaled phase window.
- 886 completed records in the full pilot exposure dataset.
- 656 completed records in the restricted pilot analytical subset.
- 8,505 completed records in the scaled 2025 data.
- 9,796 assigned records in the scaled 2025 data.

Inference analysis was done on the full pilot exposure data set, i.e., 886 records.

4.5 Measures

Three measures were defined based on the data obtained from the Safety Culture platform.

- Adoption

Completion Rate = Completed Assignments / Assigned Assignments

- Learning Effectiveness

Average assessment score of completed assignments

Pilot (full exposure) - $M = 80.10$, $SD = 21.55$, Variance = 464.30

-Scaled: $M = 83.38$, $SD = 23.14$, Variance = 535.46

- Learning Efficiency: Mean time-on-task (minutes) for completed assignments only.

Pilot: $M = 10.30$, $SD = 11.53$, Variance = 132.95

Scaled: $M = 5.98$, $SD = 7.82$, Variance = 61.15

Oversight Construct Clarification: Human oversight was incorporated into the governance structure but was not measured as a quantitative variable. Therefore, oversight was treated as a contextual condition of governance rather than a statistically controlled moderator.

4.6 Analytical Approach (Upgraded)

Descriptive Analysis (mean, standard deviation, variance, minimum, maximum) were run separately per phase. Inferential Tests were run for the the analytical rigor of the study,

- Completion Rate Difference: Two-proportion z-test

$z = 11.52$, $p < .001$

- Difference = -13.18 percentage points

95% CI = [-13.85, -12.51]

Odds Ratio (scaled vs pilot): OR = 0.16

- Assessment Score Difference: Welch's independent samples t-test

$t(1108.70) = 4.29$, $p < .001$

- Mean difference = 3.28

95% CI [1.78, 4.79]

- Cohen's $d = 0.14$ (small effect)

Time-on-Task Difference: Welch's t-test

$t(971.64) = -10.88$, $p < .001$

- Mean difference = -4.31 minutes

95% CI [-5.09, -3.53]

- Cohen's $d = -0.52$ (moderate effect)

- Regression Specification (Robustness): Linear regression models were run as a further robustness test

$Outcome_i = \beta_0 + \beta_1 DeploymentMaturity_i + \varepsilon_i$

Deployment Maturity = 0 (pilot), 1 (scale)

Results were consistent with t-test results.

4.7 Control Variables and Organizational Context

The observational data lacked complete, structured data for variables such as tenure, role type, and outlet maturity for both phases. As a result, these variables could not be modeled as covariates. However, the following can be said:
- Core training content is constant.
- Micro-agent configuration is constant.
- There is a gradual expansion throughout 2025, rather than a sharp change.

For future research, structured employee controls such as tenure, role type, and outlet age should be included.

4.8 Survivorship and Cohort Considerations

The data does not track a panel of identical individuals across both phases, as there is a change in workforce composition due to organizational expansion and attrition.

As a result:
- Longitudinal effects exist at the organizational level.
- Results should not be interpreted as a causal process for individuals.
- Exposure-adjusted active definitions reduce, but do not eliminate, effects of churn.

4.9 Causal Inference Limitations

For a non-experimental design:
- There is no randomization.
- There is no counterfactual comparison group.
- There is organizational expansion with deployment scale.

As a result, findings should be interpreted as statistically significant longitudinal associations, not causal effects of deploying AI-enabled training.

5. Results

In this section, the results from the 2024 pilot and 2025 scaled deployment of the AI-powered training micro-agents at Toscano are presented. All results are based on SafetyCulture learning trace data and are reported at the assignment record level. This approach provides a better representation of systemic behavioral patterns rather than individual-level inference.

5.1 Study Cohorts and Training Volume

The 2024 pilot phase (active users September-November) comprised 65 employees with 656 assigned course records. All assigned course records were completed. For inferential analysis and hypothesis testing, the total pilot exposure dataset with 886 completed course records was used.

The 2025 scaled phase comprised 250 active employees with 9,796 assigned course records. Out of those, 8,505 course records were completed. In both phases, the training volume and workforce engagement have been substantially increased with the outlet growth.

5.2 Adoption: Completion Rates under Scale

Adoption was measured as course completion rate.
- 2024 Pilot Phase: 100.00%

- 2025 Scaled Phase: 86.82%
- Difference: -13.18 percentage points

The z-test for two proportions revealed that the observed difference between the 2024 and 2025 phases was statistically significant.
$z = 11.52$, $p < .001$
95% CI for the difference:
[-13.85, -12.51]
Odds Ratio (2025 scaled phase and 2024 Pilot Phase):
OR = 0.16

The odds ratio indicates that the likelihood of course completion under scale deployment is significantly lower than that observed under the tightly supervised 2024 pilot phase. However, the substantial increase in training volume and workforce engagement indicates a normalized level under scale deployment rather than disengagement.

5.3 Learning Effectiveness: Assessment Performance

Descriptive Statistics (Full Exposure Data)
Pilot: $M = 80.10$, $SD = 21.55$
Scaled: $M = 83.38$, $SD = 23.14$
Inferential Test
Welch's t-test: $t(1108.70) = 4.29$, $p < .001$
Mean Difference
3.28
95% CI
[1.78, 4.79]
Effect Size
Cohen's $d = 0.14$
This represents a small but statistically significant improvement in assessment performance for scaled deployment.

5.4 Learning Efficiency: Time-on-Task

Learning efficiency is defined by mean time-on-task per assignment completed.
Descriptive Statistics
Pilot: $M = 10.30$ minutes, $SD = 11.53$
Scaled: $M = 5.98$ minutes, $SD = 7.82$
Inferential Test
Welch's t-test: $t(971.64) = -10.88$, $p < .001$
Mean Difference
-4.31 minutes
95% CI
[-5.09, -3.53]
Effect Size
Cohen's $d = -0.52$
This represents a moderate level of improvement in efficiency, where learners were able to complete modules in significantly less time with no decrease in assessment performance.

5.5 Robustness Checks

To test robustness, a number of robustness checks were performed.
Exclusion of High Volume Outlets
Exclusion of assignments from the top 10% of outlets by volume. Results for assessment performance and time-on-task maintained their direction and significance.
Stratification by Role
For those outlets where role-level identifiers were accessible, similar effects were found.

5.6 Alternative Explanations

Learning Curve / Familiarity Effect: The efficiency gain may also be partly attributed to increased familiarity with the platform rather than the effectiveness of micro-agents. Nevertheless, increased assessment performance and concomitant decreases in time-on-task indicate mastery effects.

Organizational Expansion Effects: Toscano expanded its outlets from 35 to 50 during the study duration. Organizational expansion may have brought about variability in new users. Nevertheless, consistent efficiency gain in both conditions suggests positive adaptation.

Reduced Novelty / Hawthorne Effects: The novelty effects of the pilot condition may have contributed to increased learning behavior. Nonetheless, sustained high completion rates (86.82%) and increased performance indicate that learning behavior persisted beyond novelty effects.

The interpretations presented are consistent with non-experimental design and are conceptualized as longitudinal effects rather than causality.

5.7 Summary of Longitudinal Outcomes

In summing up the results of the analysis, it is clear that there is a consistent longitudinal pattern in the data collected from the use of the AI-enabled learning system from its pilot use to its large-scale use:

- Completion rates returned to normal levels in large-scale use and remained high.
- There were increased learning effectiveness levels, evidenced by statistically significant learning effectiveness in assessment performance ($d = 0.14$).
- There were increased learning efficiencies in large-scale use, evidenced by moderate effect sizes in reductions in time-on-task ($d = -0.52$).

The results show that large-scale use of learning systems is statistically associated with sustained use and learning efficiency.

6. Discussion

This study examined how AI-enabled training micro-agents are associated with changes in adoption, learning efficiency, and oversight dynamics as deployment transitioned from a 2024 supervised pilot to a 2025 scaled organizational implementation. The longitudinal findings offer insight into how constrained, workflow-embedded AI systems function once novelty subsides, participation expands, and operational complexity increases.

Importantly, given the observational design, the findings are interpreted as statistically significant longitudinal associations rather than definitive causal effects.

6.1 From Pilot Optimality to Scaled Normalization

One of the most salient findings is the statistically significant decline in completion rates from 100% during the pilot to 86.82% under scaled deployment ($z = 11.52$, $p < .001$). While this reduction may appear negative when viewed cross-sectionally, a longitudinal interpretation suggests a more nuanced dynamic.

The pilot phase operated under tightly controlled conditions: limited cohort size, heightened managerial visibility, and structured enforcement. Under such conditions, perfect completion rates are not uncommon. As the system scaled alongside Toscano's expansion from 35 to 50 outlets, training volume increased substantially and direct supervisory enforcement attenuated.

The observed decline therefore appears consistent with normalization under operational scale rather than disengagement. The persistence of a high completion rate under dramatically increased volume suggests continued adoption of the training system under less constrained conditions. This pattern aligns with technology acceptance research emphasizing sustained use over initial compliance (Venkatesh et al., 2003).

6.2 Learning Efficiency as a Longitudinal Association

The most theoretically significant finding is the simultaneous improvement in learning effectiveness and efficiency. Assessment scores increased modestly (Cohen's $d = 0.14$), while time-on-task declined with a moderate effect size ($d = -0.52$).

The magnitude difference is noteworthy. The improvement in performance was statistically significant but small in effect size, whereas the reduction in time-on-task reflects a moderate behavioral shift. This pattern suggests that the primary longitudinal change was efficiency enhancement rather than dramatic knowledge gain.

Such an association is consistent with learning fluency effects. As exposure deepens, learners may require less cognitive overhead to navigate content while maintaining or slightly improving mastery. From a training transfer perspective, repeated reinforcement and structured sequencing may facilitate procedural fluency rather than raw score acceleration (Baldwin & Ford, 1988).

Mechanisms: Which Micro-Agent Functions Likely Contributed?Based on the typology outlined in Section 2, several micro-agent mechanisms are plausibly associated with observed efficiency gains:

• Sequencing agents likely reduced search and navigation friction by structuring progression paths.
• Nudging agents may have minimized procrastination and clustered learning into shorter, focused sessions.
• Reinforcement agents provided rapid feedback loops, enabling corrective adjustments without repeated trial-and-error.
• Friction-reduction agents likely lowered task-switching costs within the platform interface.

Importantly, these mechanisms operate at the behavioral workflow level rather than at the knowledge-content level. The moderate efficiency effect size suggests that micro-agents were associated more strongly with how learning occurred than with how much knowledge was ultimately retained.

This interpretation remains consistent with human–AI complementarity theory (Fügener et al., 2019; Hemmer et al., 2025), where AI systems scaffold processes while humans retain judgment authority.

6.3 Trust Calibration and the Absence of Algorithm Aversion

The absence of performance deterioration or completion collapse over time suggests that algorithm aversion effects did not dominate behavior (Dietvorst et al., 2015). Continued engagement and modest performance gains indicate that user trust may have been calibrated rather than eroded.

Several contextual features may have contributed to this outcome:

1. Micro-agents did not exercise evaluative authority.
2. Managers retained interpretive and enforcement control.
3. The system generated nudges rather than binding directives.

These characteristics align with established human–AI design guidelines emphasizing transparency, controllability, and appropriate reliance (Amershi et al., 2019). By constraining algorithmic authority, the deployment minimized conditions under which aversion or resistance might emerge.

However, it is important to note that trust was not directly measured in this study. The interpretation of calibrated trust is therefore inferential and should be treated cautiously.

6.4 Human Oversight as Contextual Governance

The results suggest that human oversight functioned as a stabilizing governance condition rather than as a statistically modeled moderator. Across both deployment phases, managers retained responsibility for monitoring progress, addressing exceptions, and contextualizing performance analytics.

While oversight intensity differed structurally between pilot and scale, it was not directly quantified. Therefore, claims regarding governance effects remain conceptual rather than empirically estimated.

That said, the coexistence of efficiency gains and stable assessment performance indicates that oversight mechanisms likely prevented superficial acceleration or metric gaming. The presence of managerial accountability may have mitigated risks associated with over-automation, aligning with responsible AI governance principles (Papagiannidis et al., 2025).

Future research should operationalize oversight intensity directly, incorporating escalation frequency, managerial intervention rates, or audit logs as measurable indicators.

6.5 Alternative Interpretations

Several alternative explanations warrant consideration.

Learning Curve Effects, Efficiency gains may reflect increased familiarity with the platform rather than micro-agent influence per se. However, the persistence of improved assessment performance suggests that reductions in time-on-task were not solely attributable to navigation fluency.

Organizational Expansion, The concurrent expansion from 35 to 50 outlets introduced workforce heterogeneity. Such expansion typically increases variability in onboarding quality. The absence of performance deterioration under expansion conditions suggests system resilience, though this cannot be causally attributed.

Novelty and Hawthorne Effects, The pilot phase likely benefited from novelty and heightened attention. The maintenance of high completion rates and improved efficiency beyond the pilot window suggests that effects were not purely novelty-driven.

These alternative interpretations reinforce the importance of cautious inference.

6.6 Theoretical Implications

The findings extend technology acceptance literature by demonstrating that deployment maturity is associated with normalization in adoption behavior and efficiency gains over time (Davis, 1989; Venkatesh & Davis, 2000). Pilot outcomes alone are insufficient predictors of scaled performance.

The study also contributes to training research by providing objective behavioral evidence of longitudinal efficiency effects, addressing measurement concerns raised in prior work (Baldwin & Ford, 1988; Podsakoff et al., 2003).

Finally, by conceptualizing AI-enabled training systems as constrained micro-agents embedded within governance structures, the study advances human–AI scholarship beyond automation-versus-autonomy dichotomies. The evidence suggests that bounded, workflow-level agency may be associated with sustainable performance patterns.

6.7 Practical Implications

For practitioners, the results indicate that scaled normalization should be anticipated rather than resisted. Perfect pilot completion rates are unlikely to persist under expansion conditions.

More importantly, the moderate efficiency gains suggest that AI-enabled micro-agents may support operational scalability by reducing time costs per learning unit. However, these systems should be implemented alongside explicit governance mechanisms.

As efficiency increases, oversight becomes more—not less—important. Organizations should monitor not only completion but also score stability and anomaly patterns to ensure that acceleration does not compromise learning integrity.

7. Limitations and Future Research

This research provides a longitudinal, behavioral-based approach to AI-supported training of micro-agents within an organizational context. Nevertheless, there are limitations to this research, and these limitations provide a framework for future research.

7.1 Contextual and Generalizability Limitations

One of the first limitations is that, despite its focus on a rapidly expanding Italian casual dining restaurant chain, the research is specific to a particular context. Although Toscano is representative of a common frontline-intensive service context, there is a possibility that the results of this research cannot be generalized across other industry contexts, especially those with dissimilar workforce structures, regulatory requirements, and learning environments.

In addition, the fact that the organization expanded from 35 outlets to 50 outlets during the course of the research is a contextual constraint. The impact of expansion can influence learning outcomes. Although the longitudinal approach is representative of real-world dynamics, there is a possibility that expansion can influence learning outcomes. As such, future research should focus on using multiple organizations and include contextual growth variables as moderators.

7.2 Absence of Experimental Control and Limitations in Causal Inference

Secondly, the current study employs an observational rather than an experimental or quasi-experimental research framework. Although this improves ecological validity, causal attribution is limited. The observed improvements in learning efficiency and assessment outcomes are described as statistically significant rather than causal associations of AI-based micro-agents with learning efficiency and assessment outcomes.

Potential confounding variables:

- Changes in workforce composition
- Variability in tenure
- Variability in management enforcement
- Growth of organizational units

Future studies could use quasi-experimental methods, staggered rollouts, or feature toggling of micro-agents to isolate specific behavioral mechanisms.

7.3 Hawthorne and Novelty Effects

Thirdly, the 2024 pilot phase may have experienced Hawthorne or novelty effect bias, where employees participating in the doctoral study may have recorded higher engagement due to raised visibility and possibly evaluation of their behavior or novelty of the digital tool.

Though sustained improvements in efficiency observed during the 2025 phase minimize novelty effect bias, as observed in the current study, it is not possible to rule out novelty effect bias entirely as observed during the 2024 phase.

Future studies could use baseline or delayed treatment groups to help disentangle novelty effect bias from behavioral adaptation.

7.4 Platform Dependency and System-Specific Effects

Fourth, it is important to note that this study was based entirely on the SafetyCulture AI-based training platform. While this platform represents a constrained, workflow-based AI micro-agent, it is possible that platform-specific effects may have contributed to observed effects. Future research should consider comparing multiple AI-based training systems in order to better understand the effects of micro-agent design, transparency, configuration, and autonomy levels on adoption trajectories and efficiency gains.

7.5 Measurement and Operational Definitions

There are several methodological issues that require clarification.

Active Employee Definition: Active employees were defined based on those individuals who had shown continued learning engagement within each phase of study. Specifically, it was defined as those individuals with at least one course completion and platform interaction within each phase. This was done in order to reduce attrition bias and survivor bias. However, it is possible that other definitions, based on thresholds of minimum exposure or tenure, could have produced different results. Future research should consider sensitivity testing across different thresholds.

Record-Level vs. Employee-Level Analysis: While analysis was performed at the record level, descriptive statistics were provided at the employee level. Completion rate and calculation of time-on-task represent aggregated course-level data. Future research should consider using a multilevel modeling structure, nesting course records within employees, and employees within outlets.

7.6 Distributional Assumptions and Statistical Scope

Although inferential tests were performed, distribution tests such as skewness, kurtosis, and normality tests were not included in the regression model framework. Future research should include:

- Normality tests (e.g., Shapiro–Wilk)
- Log transformations for time variables
- Heteroskedasticity tests
- Robustness tests for non-parametric tests

7.7 Temporal Scope and System Maturation

Although the current research covers pilot and scaled phases over two years, it does not cover long-term post-adoption system maturation beyond the scaled phase. Information systems research has emphasized the importance of post-adoption and routinization behaviors, which take place over long periods and can include effects such as habituation, oversight recalibration, efficiency plateauing, and even engagement decay. Long-term longitudinal research would be useful for studying:

- Habituation
- Oversight recalibration
- Efficiency plateauing
- Engagement decay

These are interesting areas for future research on post-adoption information systems maturation.

7.8 Future Research Agenda

Based on the limitations of the current research, future research should be conducted in the following four directions:

1. Experimental Micro-Agent Manipulation

Randomly activate and deactivate particular micro-agent features, for example, nudging, sequencing, and reinforcement.

2. Governance Intensity Modeling

Quantify oversight and model it in the regression model framework for moderation effect testing.

3. Contextual Growth Moderation

Model organizational expansion variables in the regression model framework for moderation effect testing.

4. Cross-Platform Comparative Research

Explore the effect of differences in AI-enabled training architectures on longitudinal adoption and efficiency patterns.

8. Conclusion

However, as artificial intelligence becomes more integrated into organizational learning systems, it is also expected that the evaluation process should move beyond pilot-based snapshots and into longitudinal data based on actual operating conditions. This research paper makes a contribution towards this progression by offering behavior-based longitudinal evidence for AI-enabled training micro-agents in an organization moving from pilot to scale in the context of a frontline-intensive organization. The results indicate that, in terms of moving from pilot to scale, there was:

- Improved completion rates
- Improved learning efficiency with moderate effect size
- Improved assessment results
- Stable adoption even in the context of organizational expansion

These results suggest that AI micro-agents in a constrained state within human-governed systems are related to sustained behavioral adaptation over time. However, it should be noted that the research paper does not suggest that the efficiency gains were in any way related to the AI micro-agents themselves. Rather, it suggests that in the context of a human-in-the-loop governance structure, AI-enabled training was related to improved fluency and mastery at scale.

From a conceptual perspective, this research offers a significant contribution in terms of advancing a micro agent-based understanding of AI-supported training, which is task-bound, workflow-integrated, and behavioral in nature, shaping learning processes without replacing human agency. This understanding offers a middle ground in addressing technological optimism in automation and skepticism in governance. From a theoretical perspective, this research offers a significant contribution in terms of advancing:

- Post-adoption IS maturation theory
- Training transfer and reinforcement theory
- Human-AI complementarity theory

From a methodological perspective, this research offers a significant contribution in terms of utilizing learning trace data provided by the platform, which offers objective insights in understanding longitudinal effects, eliminating self-reported bias. Practically, this research offers a significant contribution in terms of understanding that AI-supported training systems should be built with normalization in mind, not perfection, in terms of pilot effects. Governance structures are important in achieving increased efficiency. In conclusion, it can be said that AI-supported training systems hold significant promise not in terms of their intelligence, per se, but in terms of responsibly constraining micro-agents within accountable human systems. By moving away from understanding pilot effects, this research offers a replicable evaluation model for organizations looking to incorporate AI in their training systems.

References

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., … Horvitz, E. (2019). *Guidelines for Human-AI Interaction*. **CHI 2019**. https://doi.org/10.1145/3290605.3300233

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology, 41*(1), 63–105. https://doi.org/10.1111/j.1744-6570.1988.tb00632.x

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

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126. https://doi.org/10.1037/xge0000033

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2019). Cognitive challenges in human–AI collaboration (working paper). SSRN. https://doi.org/10.2139/ssrn.3368813

Hemmer, P., Schemmer, M., Kühl, N., Vössing, M., & Satzger, G. (2025). Complementarity in human–AI collaboration. *European Journal of Information Systems*. https://doi.org/10.1080/0960085X.2025.2475962

Hosen, S., Hamzah, S. R., Ismail, I. A., Alias, S. N., Abd Aziz, M. F., & Rahman, M. M. (2024). Training & development, career development, and organizational commitment as the predictor of work performance. *Heliyon, 10*(1), e23903. https://doi.org/10.1016/j.heliyon.2023.e23903

Madanchian, M., Taherdoost, H., & Mohamed, N. (2023). AI-based human resource management tools and techniques: A systematic literature review. *Procedia Computer Science, 229*, 367–377. https://doi.org/10.1016/j.procs.2023.12.039

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). *Responsible artificial intelligence governance: A review and research framework.* Journal of Strategic Information Systems, 34, Article 101885. https://doi.org/10.1016/j.jsis.2024.101885

Pappas, I. O., Giannakos, M. N., & Sampson, D. G. (2019). Fuzzy set analysis for learning systems: The role of complex concepts and human factors. *Computers in Human Behavior, 92*, 646–659. https://doi.org/10.1016/j.chb.2018.03.032

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research. *Journal of Applied Psychology, 88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model. *Management Science, 46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

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