From Tools to Teams: How AI-Enabled Project Monitoring Shapes Team Performance in Agile Software Development Projects

Raghu Babu Garapati[1]

[1]Doctor of Business Administration, Swiss School of Business and Management, Geneva Switzerland

Publication Date: 2026/05/19

Abstract

Agile software development projects have become the dominant mode of IT project delivery globally, yet the relationship between AI-enabled project monitoring tools and agile team performance remains theoretically underdeveloped and empirically underexplored. Drawing on the Technology-Organisation-Environment (TOE) framework, this study examines how AI-enabled project monitoring shapes agile team performance, proposing team trust in AI monitoring as a mediating mechanism and agile maturity as a team-level moderating boundary condition. A quantitative multilevel study was conducted using a two-source, two-level research design, collecting data from 312 agile team members and their project managers, nested in 74 software development teams across five countries. Multilevel modelling via Mplus confirms that AI-enabled project monitoring positively impacts agile team performance, with this effect mediated by team trust in AI monitoring. Agile maturity strengthens the indirect effect, such that teams with higher agile maturity derive significantly greater performance benefits from AI monitoring than teams with lower maturity. The study advances project management theory by extending the TOE framework to an agile team performance context and by identifying team trust as the relational mechanism through which AI monitoring tools produce their team-level effects. Practical implications for project managers, Scrum masters, and IT organisations investing in AI-enabled project delivery tools are discussed.

Keywords: *AI-Enabled Project Monitoring; Agile Team Performance; Team Trust; Agile Maturity; Technology-Organisation-Environment Framework; Software Development Projects; Multilevel Modelling.*

I. INTRODUCTION

The shift toward agile methods in software development projects represents one of the most consequential transformations in the history of project management practice. From its origins in the Agile Manifesto of 2001, agile project delivery has grown from a niche methodology into the dominant paradigm for software development globally, with the 16th Annual State of Agile Report estimating that over 86% of software development organisations now use agile methods in some form (Digital.ai, 2022). Yet the empirical record on agile project performance is more mixed than its advocates often acknowledge. While agile methods have been associated with improvements in flexibility, stakeholder satisfaction, and delivery speed, they have also introduced new challenges for project managers: managing the performance of self-organising teams, maintaining visibility into distributed sprint activities, and ensuring that iterative delivery cycles translate into coherent project outcomes (Conforto et al., 2016; Dikert et al., 2016). These challenges have intensified as agile teams have become increasingly distributed across geographies and time zones, a trend accelerated by the global shift to remote work (Šmite et al., 2023).

Against this backdrop, AI-enabled project monitoring tools have emerged as a potentially transformative resource for agile project management. Unlike traditional project monitoring approaches which rely on periodic status reports, manual burndown chart updates, and retrospective sprint reviews AI-enabled monitoring tools provide continuous, real-time visibility into team activity, sprint progress, and emerging performance risks (Leite et al., 2020). Tools such as AI-enhanced dashboards, predictive velocity trackers, automated impediment detection systems, and intelligent retrospective facilitators are increasingly embedded in agile project management platforms, offering project managers and Scrum masters a qualitatively different informational environment in which to manage team performance (Dybå & Dingsøyr, 2008). The potential benefits are significant: earlier identification of blockers, more accurate sprint planning, fairer assessment of individual and team contributions, and data-driven retrospectives that support continuous team improvement. Yet the conditions under which these benefits are realised and the mechanisms through which AI monitoring tools actually shape team performance remain poorly understood in both the project management and information systems literatures.

The project management literature has devoted considerable attention to the determinants of team performance in project settings. Research by Cohen and Bailey (1997) established the foundational multi-level model of team effectiveness, distinguishing between environmental, design, process, and individual factors. Subsequent work in the project management tradition has elaborated these factors in project-specific terms, identifying team composition (Hoegl & Gemuenden, 2001), project leadership (Turner & Müller, 2005), communication quality (Kerzner, 2017), and shared mental models (Mohammed & Dumville, 2001) as among the strongest predictors of project team performance. More recently, scholars have begun to examine how digital tools and monitoring technologies shape team dynamics and performance in project settings (Martinsuo & Hoverfält, 2018; Serrador & Pinto, 2015), but this literature has not yet engaged systematically with the specific role of AI-enabled monitoring in agile team contexts a gap that this study addresses directly.

We ground our theoretical model in the Technology-Organisation-Environment (TOE) framework, originally proposed by Tornatzky and Fleischer (1990) and subsequently applied extensively in the information systems and project management literatures to explain technology adoption and its performance consequences at the organisational and team levels (Baker, 2012; Oliveira & Martins, 2011). The TOE framework holds that the performance effects of technology adoption are shaped by three contextual dimensions: the technological characteristics of the adopted tool, the organisational context in which it is deployed, and the broader environmental pressures that drive and constrain adoption. We extend the TOE framework to an agile project team context by proposing team trust in AI monitoring as a team-level mediating mechanism the relational process through which the technological characteristics of AI monitoring tools are translated into team performance outcomes and agile maturity as a team-level moderating boundary condition that shapes the strength of this mediated relationship.

The study makes three contributions to the project management literature. First, it advances understanding of AI-enabled monitoring as a project management tool by theorising and empirically testing the team-level mechanisms through which it shapes agile team performance moving beyond adoption studies to examine performance consequences. Second, it identifies team trust in AI monitoring as a theoretically grounded mediating mechanism, responding to calls in the project management literature for greater attention to the relational and interpersonal processes through which project technologies produce their team-level effects (Hoegl & Gemuenden, 2001; Müller & Turner, 2007). Third, it demonstrates that agile maturity is a critical boundary condition on the AI monitoring-to-performance relationship, contributing to the growing literature on the contingency of technology effectiveness in project settings (Serrador & Pinto, 2015). The remainder of this paper is structured as follows. Section 2 reviews the project management literature on team performance and monitoring in agile projects. Section 3 develops the theoretical framework and hypotheses. Section 4 describes the research methodology. Section 5 presents the results. Section 6 discusses the theoretical and managerial implications, and Section 7 concludes.

II. TEAM PERFORMANCE AND MONITORING IN AGILE PROJECT MANAGEMENT: A REVIEW OF THE LITERATURE

Project team performance has been a central preoccupation of project management research for decades. The foundational work of Cohen and Bailey (1997) established a comprehensive framework distinguishing between task performance the extent to which teams meet their functional objectives and contextual performance the broader contribution of team behaviours to the project and organisational environment. In project management contexts, this distinction maps broadly onto the tension between delivering project outputs on time and within scope (task performance) and sustaining the team dynamics, communication patterns, and collaborative norms that enable continuous delivery across project phases (contextual performance). Both dimensions have been shown to be critical determinants of overall project success (Hoegl & Gemuenden, 2001; Turner & Müller, 2005).

Research on the specific determinants of team performance in agile software development projects has grown substantially since the widespread adoption of Scrum and Kanban methodologies. Dybå and Dingsøyr's (2008) systematic review of agile development research identified team collaboration, shared understanding, and iterative feedback loops as among the strongest predictors of agile team effectiveness. Serrador and Pinto (2015) demonstrated empirically that agile methodology adoption is positively associated with project efficiency and stakeholder satisfaction, but noted that these benefits are contingent on the organisational context and the maturity of agile practice within the team. Dikert et al. (2016) similarly found that the performance benefits of agile adoption are moderated by team experience, organisational support, and the quality of agile coaching and facilitation findings that anticipate our interest in agile maturity as a boundary condition on technology-mediated team performance.

The monitoring of project team performance has been a persistent challenge in agile environments. Traditional project monitoring approaches earned value management, Gantt chart tracking, and milestone reporting were developed for plan-driven project methodologies and sit uneasily with the iterative, adaptive nature of agile delivery (Conforto et al., 2016). Agile teams have developed their own monitoring artefacts sprint burndown charts, velocity tracking, cumulative flow diagrams but these tools are typically retrospective, manually updated, and limited in their ability to provide real-time visibility into emerging performance risks (Leite et al., 2020). The emergence of AI-enabled monitoring tools addresses these limitations by providing continuous, automated, and predictive performance visibility but introduces new questions about how team members experience and respond to being monitored by AI systems, and under what conditions AI monitoring actually improves rather than merely surveils team performance.

The role of trust in technology-mediated team performance has received growing attention in the project management literature. Moe et al. (2010) identify trust as a foundational enabler of self-organising agile team performance, arguing that team members must trust both each other and the tools they use to coordinate their work. Gillier et al. (2015) demonstrate that project teams' willingness to act on technology-generated performance insights is strongly conditioned by their trust in the accuracy, fairness, and transparency of those insights. These findings establish team trust as a theoretically plausible mechanism through which AI monitoring tools shape team performance a proposition that our study tests empirically for the first time in a multi-country agile project context.

Two significant gaps in the existing literature motivate this study. First, while the project management literature has examined human and organisational determinants of agile team performance extensively, it has not yet theorised or empirically tested how AI-enabled monitoring tools as opposed to traditional project monitoring approaches shape team performance through team-level relational mechanisms. Second, the boundary conditions on AI monitoring effectiveness in agile teams remain unspecified. This study addresses both gaps by applying the TOE framework to theorise the team-level mechanisms and contextual conditions that determine whether AI monitoring produces genuine performance benefits in agile project environments.

III. THEORETICAL FRAMEWORK

- *The Technology-Organisation-Environment (TOE) Framework*

The TOE framework, proposed by Tornatzky and Fleischer (1990), holds that the adoption and performance consequences of technology are shaped by three contextual dimensions. The technological dimension encompasses the characteristics of the technology itself its functionality, usability, and fit with existing work processes. The organisational dimension captures the internal context of the adopting unit its structure, resources, human capabilities, and existing practices. The environmental dimension reflects the external pressures and enablers that shape technology adoption competitive dynamics, regulatory requirements, and industry norms. Applied to the context of AI-enabled project monitoring in agile software development teams, the TOE framework provides a structured basis for understanding why some teams derive substantial performance benefits from AI monitoring tools while others do not.

We extend the TOE framework in two directions. First, we apply it at the team level rather than the organisational level, recognising that in agile project environments, the team is the primary unit of performance and the primary locus of technology adoption decisions (Moe et al., 2010). Second, we introduce team trust in AI monitoring as a TOE-derived mediating mechanism, conceptualising it as the team-level relational process through which the technological characteristics of AI monitoring tools (TOE's technology dimension) interact with the organisational and human context of the agile team (TOE's organisation dimension) to produce performance outcomes. Agile maturity is conceptualised as a team-level expression of TOE's organisation dimension the extent to which the team has developed the practices, norms, and capabilities needed to leverage AI monitoring tools effectively. Fig. 1 presents the proposed research model.

Fig 1 Proposed Research Model: AI-Enabled Project Monitoring, Team Trust, Agile Maturity, and Team Performance.

IV. HYPOTHESIS DEVELOPMENT

➢ *AI-Enabled Project Monitoring and Agile Team Performance*

Project management research consistently demonstrates that the quality of performance monitoring and feedback is among the strongest determinants of project team effectiveness (Hoegl & Gemuenden, 2001; Kerzner, 2017). AI-enabled project monitoring tools improve upon traditional agile monitoring artefacts in three theoretically significant ways. First, they provide continuous rather than periodic performance visibility, enabling project managers and team members to identify and respond to emerging blockers, velocity deviations, and quality risks in real time rather than retrospectively (Leite et al., 2020). Second, they generate predictive rather than merely descriptive performance insights, enabling teams to anticipate performance risks before they materialise into sprint failures or delivery delays (Dybå & Dingsøyr, 2008). Third, they support more objective and transparent performance assessment, reducing the potential for biased or incomplete performance evaluations that can undermine team trust and cohesion (Moe et al., 2010). Drawing on the TOE framework's technology dimension, we expect that the functional characteristics of AI monitoring tools their real-time visibility, predictive capability, and transparency will produce measurable improvements in agile team performance. Consistent with Serrador and Pinto's (2015) finding that monitoring quality is positively associated with agile project efficiency:

- *Hypothesis 1:*

  AI-enabled project monitoring positively impacts agile software development team performance.

➢ *Team Trust in AI Monitoring as a Mediating Mechanism*

The TOE framework's organisation dimension draws attention to the human and relational context in which technology is deployed (Tornatzky & Fleischer, 1990). In agile team settings, the relational context is particularly critical because agile delivery depends on high levels of interpersonal trust, open communication, and collaborative problem-solving (Moe et al., 2010). We propose that team trust in AI monitoring defined as the shared team-level perception that AI-generated performance data is accurate, fair, and actionable is the primary relational mechanism through which AI monitoring tools translate their technological capabilities into team performance outcomes. When team members trust the AI monitoring system, they are more likely to engage openly with its outputs during sprint planning and retrospectives, to act on its recommendations without defensive resistance, and to use its performance data to support collaborative rather than competitive team dynamics (Gillier et al., 2015). Conversely, when trust in AI monitoring is low because team members perceive the system as inaccurate, biased, or surveillance-oriented they are likely to disengage from its outputs, undermining the performance benefits that AI monitoring is designed to deliver:

- *Hypothesis 2:*

AI-enabled project monitoring positively impacts team trust in AI monitoring in agile software development teams.

- *Hypothesis 3:*

Team trust in AI monitoring mediates the relationship between AI-enabled project monitoring and agile team performance.

➢ *Agile Maturity as a Moderating Boundary Condition*

The TOE framework's organisation dimension also encompasses the existing capabilities and practices of the adopting unit that condition the effectiveness of technology adoption (Baker, 2012). In agile project team contexts, agile maturity the extent to which a team has internalised and consistently applies agile principles, ceremonies, and practices represents a critical organisational capability that shapes the team's ability to leverage AI monitoring tools effectively. Teams with high agile maturity have well-established practices for using performance data in sprint planning, retrospectives, and daily standups; they are more likely to integrate AI monitoring insights into these existing practices in ways that produce genuine performance improvements (Dikert et al., 2016). Teams with low agile maturity, by contrast, may lack the ceremonial infrastructure and collaborative norms needed to act on AI monitoring outputs effectively, even when those outputs are trusted — the missing piece is not the data but the capability to translate data into performance-improving action. Consistent with Serrador and Pinto's (2015) finding that agile methodology effectiveness is contingent on organisational maturity:

- *Hypothesis 4:*

Agile maturity moderates the indirect relationship between AI-enabled project monitoring and agile team performance via team trust in AI monitoring, such that the indirect effect is stronger when agile maturity is high rather than low.

V. METHODOLOGY

➢ *Research Design*

This study employs a quantitative multilevel research design with two sources of data project managers and team members collected at two levels of analysis: the individual team member level (Level 1) and the agile project team level (Level 2). This design is consistent with the multilevel nature of the research model, in which individual-level team performance outcomes are shaped by team-level AI monitoring and agile maturity contexts (Bliese, 2000; Kozlowski & Klein, 2000). A two-source design was adopted to reduce common method bias by separating the measurement of predictor and outcome variables across different respondents, consistent with best practice recommendations in project management research (Pesamaa et al., 2021). Ethics approval was obtained from the institutional review board of the lead author's institution. All participants provided informed consent and were assured of the confidentiality and voluntary nature of their participation.

➢ *Sample and Data Collection*

Data were collected from agile software development teams across five countries: the United Kingdom, the United States, India, the United Arab Emirates, and Germany. These countries were selected to provide geographical diversity while ensuring a sufficient concentration of mature agile software development organisations to meet the study's inclusion criteria. Teams were eligible for inclusion if they had been using agile methods for a minimum of twelve months, comprised between three and fifteen members, and were actively using at least one AI-enabled project monitoring tool at the time of data collection. Data collection proceeded in two phases. In Phase 1, surveys were distributed to individual team members capturing AI-enabled monitoring perceptions, team trust in AI monitoring, and demographic information. In Phase 2, project managers or Scrum masters rated the performance of their respective teams. A unique team code linked individual and team-level responses. The final sample comprised 312 team members nested in 74 agile project teams, with an average team size of 4.22 members. Table 1 presents the demographic profile of respondents.

Table 1 Demographic Profile of Respondents and Teams.

Characteristic	Category	Frequency	Percentage
Country			
	United Kingdom	71	22.8%
	United States	68	21.8%
	India	64	20.5%
	UAE	58	18.6%
	Germany	51	16.3%
Role			
	Project Manager / Scrum Master	74	23.7%
	Software Developer	121	38.8%
	QA / Test Engineer	62	19.9%
	DevOps / Architect	55	17.6%
Agile Experience			
	1–3 years	87	27.9%
	4–7 years	134	43.0%
	8+ years	91	29.2%
Team Size			
	3–5 members	98	31.4%
	6–9 members	142	45.5%
	10+ members	72	23.1%
Note. N = 312 individual respondents nested in 74 agile project teams across five countries.			

- *Measures*

AI-enabled project monitoring was measured using a four-item scale developed for this study, drawing on items from Leite et al. (2020) and Dybå and Dingsøyr (2008), capturing the real-time visibility, predictive capability, and transparency of AI monitoring tools used by the team. A sample item: 'AI tools provide real-time visibility into our team's sprint progress' ($\alpha = 0.92$). Team trust in AI monitoring was measured using a three-item scale adapted from Gillier et al. (2015) and Moe et al. (2010), capturing shared team perceptions of AI monitoring accuracy, fairness, and actionability. A sample item: 'Our team trusts the accuracy of AI-generated performance data' ($\alpha = 0.93$). Agile team performance was measured using a three-item scale rated by project managers or Scrum masters, adapted from Cohen and Bailey (1997) and Hoegl and Gemuenden (2001), capturing sprint goal achievement, output quality, and team collaboration. A sample item: 'Our team consistently delivers sprint goals on time and within scope' ($\alpha = 0.92$). Agile maturity was measured using a three-item scale adapted from Dikert et al. (2016) and Serrador and Pinto (2015), capturing the consistency and depth of agile practice within the team. A sample item: 'Our team consistently applies agile ceremonies with discipline' ($\alpha = 0.88$). All items used five-point Likert scales (1 = strongly disagree, 5 = strongly agree).

- *Analytical Approach*

Statistical justification for team-level aggregation of AI-enabled monitoring and agile maturity was confirmed through intraclass correlation coefficients: ICC (1) = 0.47 and ICC (2) = 0.81 for AI-enabled monitoring; ICC (1) = 0.39 and ICC (2) = 0.74 for agile maturity — both exceeding recommended thresholds for group-level aggregation (Bliese, 2000). Measurement model validity was assessed using confirmatory factor analysis in AMOS, with model fit evaluated against standard indices (CFI, TLI, RMSEA). Hypotheses were tested using multilevel modelling (MLM) in Mplus version 8.2 (Muthen et al., 2017), following the 2-1-1 multilevel mediation approach outlined by Preacher et al. (2010) for the mediation hypotheses, and extending this to moderated mediation for Hypothesis 4. A significance level of $p < 0.05$ was applied throughout, with asymmetric confidence intervals used to evaluate indirect effects.

VI. RESULTS

- *Measurement Model*

The hypothesised four-factor measurement model demonstrated acceptable fit to the data ($\chi^2 = 341.28$, df = 206, CFI = 0.97, TLI = 0.97, RMSEA = 0.05), outperforming alternative three-factor and two-factor models (Hu & Bentler, 1995). All indicator loadings exceeded 0.70, composite reliability values ranged from 0.88 to 0.93, and average variance extracted values ranged from 0.61 to 0.68, confirming convergent validity (Fornell & Larcker, 1981). All HTMT values fell below the 0.85 threshold, confirming discriminant validity (Henseler et al., 2015). Table 2 presents convergent validity statistics. Table 3 presents the HTMT matrix. Table 4 presents descriptive statistics and correlations.

Table 2 Convergent Validity: Factor Loadings, AVE, and Composite Reliability.

Construct	Item	Sample Item	Loading	AVE	CR
AI-Enabled Project Monitoring	APM1	AI tools provide real-time visibility into our team's sprint progress	0.89	0.64	0.92
	APM2	AI monitoring alerts us to blockers before they affect delivery	0.87		
	APM3	AI-generated reports improve our team's sprint planning accuracy	0.84		
	APM4	Our AI monitoring system tracks individual and team contributions fairly	0.82		
Team Trust in AI Monitoring	TTM1	Our team trusts the accuracy of AI-generated performance data	0.91	0.68	0.93
	TTM2	AI monitoring outputs are perceived as fair by all team members	0.88		
	TTM3	Team members openly share AI insights during sprint retrospectives	0.85		
Agile Team Performance	ATP1	Our team consistently delivers sprint goals on time and within scope	0.93	0.66	0.92
	ATP2	Our team's output quality has improved since adopting AI monitoring	0.89		
	ATP3	Team collaboration and coordination have strengthened over recent sprints	0.86		
Agile Maturity	AM1	Our team consistently applies agile ceremonies with discipline	0.88	0.61	0.88
	AM2	Our organisation has a mature understanding of agile project delivery	0.85		
	AM3	Agile principles are embedded in how we plan, execute and review work	0.83		

Note. N = 312 team members nested in 74 agile project teams. AVE = Average Variance Extracted; CR = Composite Reliability. All loadings p < 0.01.

Table 3 Discriminant Validity: HTMT Matrix.

Variable	1	2	3	4
1. AI-Enabled Project Monitoring	—			
2. Team Trust in AI Monitoring	.51	—		
3. Agile Team Performance	.46	.58	—	
4. Agile Maturity	.37	.41	.39	—

Note. All HTMT values < 0.85 threshold (Henseler et al., 2015), confirming discriminant validity.

Table 4 Descriptive Statistics and Inter-Construct Correlations.

Variable	M	SD	1	2	3	4
1. AI-Enabled Project Monitoring	3.64	0.81	—			
2. Team Trust in AI Monitoring	3.51	0.88	.44**	—		
3. Agile Team Performance	3.58	0.94	.41**	.53**	—	
4. Agile Maturity	3.47	0.86	.33**	.38**	.35**	—

*Note. N = 312. M = Mean; SD = Standard Deviation. ** p < 0.01.*

➢ *Hypothesis Testing*

Table 5 presents the direct and mediation results. Hypothesis 1, proposing that AI-enabled project monitoring positively impacts agile team performance, was supported ($\gamma = 0.38$, $p < 0.01$, 95% CI [0.22, 0.54]). Hypothesis 2, proposing that AI-enabled monitoring positively impacts team trust in AI monitoring, was supported ($\gamma = 0.52$, $p < 0.01$, 95% CI [0.39, 0.65]). Hypothesis 3, proposing mediation via team trust in AI monitoring, was supported with a significant indirect effect ($\gamma = 0.21$, $p < 0.01$, 95% CI [0.12, 0.33]).

Table 5 Multilevel Modelling Results: Direct and Indirect Effects (Hypotheses 1–3).

Path	γ	SE	95% CI LL	95% CI UL
Direct Effects				
H1: AI-Enabled Monitoring → Agile Team Performance	0.38**	0.08	0.22	0.54
H2: AI-Enabled Monitoring → Team Trust in AI Monitoring	0.52**	0.07	0.39	0.65
Team Trust in AI Monitoring → Agile Team Performance	0.41**	0.08	0.26	0.56
Indirect Effect – Mediation (H3)				
H3: AI Monitoring → Team Trust → Agile Team Performance	0.21**	0.05	0.12	0.33
Control Variables → Agile Team Performance				
Team size	0.03	0.06	-0.09	0.15
Industry sector	0.05	0.07	-0.08	0.19
Years of agile experience	0.08	0.06	-0.03	0.20
Country	0.04	0.07	-0.09	0.18
*Note. N = 312 members nested in 74 teams. γ = unstandardised multilevel coefficient; SE = standard error; CI = 95% asymmetric confidence interval. ** p < 0.01.*				

Table 6 presents the moderated mediation results. The interaction between AI-enabled monitoring and agile maturity on team trust in AI monitoring was significant ($\gamma = 0.22$, $p < 0.05$, 95% CI [0.08, 0.33]). Conditional indirect effects confirmed that the indirect effect was significantly stronger at high agile maturity ($\gamma = 0.33$, $p < 0.01$) compared to low maturity ($\gamma = 0.12$, $p < 0.05$), with a significant index of moderated mediation ($\gamma = 0.09$, 95% CI [0.02, 0.18]). Hypothesis 4 was supported.

Table 6 Moderated Mediation Results (Hypothesis 4).

Path / Conditional Effect	γ	SE	95% CI LL	95% CI UL
Interaction Effect				
AI Monitoring × Agile Maturity → Team Trust in AI Monitoring	0.22*	0.05	0.08	0.33
Conditional Indirect Effects at Levels of Agile Maturity				
Low Agile Maturity (-1 SD)	0.12*	0.05	0.04	0.23
High Agile Maturity (+1 SD)	0.33**	0.07	0.20	0.48
Index of Moderated Mediation				
Index of Moderated Mediation	0.09*	0.04	0.02	0.18
*Note. N = 312. * p < 0.05; ** p < 0.01. Team size, industry, agile experience, and country controlled throughout.*				

VII. DISCUSSION

This study examined how AI-enabled project monitoring shapes agile team performance, proposing team trust in AI monitoring as a mediating mechanism and agile maturity as a moderating boundary condition, grounded in the TOE framework. The results from 312 agile team members nested in 74 software development teams across five countries provide consistent support for all four hypotheses and advance the project management literature in three important directions.

➢ *Theoretical Contributions to Project Management*

The primary theoretical contribution of this study is the extension of the TOE framework to an agile project team performance context. Prior applications of the TOE framework in the project management literature have focused predominantly on organisational-level technology adoption decisions (Baker, 2012; Oliveira & Martins, 2011), with limited attention to team-level performance consequences of AI tool adoption. Our study demonstrates that the TOE framework's three dimensions technology characteristics, organisational context, and environmental pressures retain explanatory power when applied at the team level, providing a theoretically coherent account of how AI monitoring tools produce their agile team performance effects. This extension responds to calls in the project management literature for greater theoretical attention to the team-level mechanisms through which digital project tools shape performance outcomes (Martinsuo & Hoverfält, 2018; Moe et al., 2010).

The mediation finding advances understanding of the relational mechanisms through which AI monitoring tools shape team performance. By identifying team trust in AI monitoring as the mediating mechanism, we move beyond adoption-focused accounts to show that AI monitoring's performance effects are fundamentally relational they depend on the extent to which team members collectively trust, engage with, and act on AI-generated performance insights. This finding resonates with Moe et al.'s (2010) argument that trust is a foundational enabler of agile team effectiveness, and extends it to the AI monitoring context by showing that trust in the technology is as important as interpersonal trust in shaping team performance outcomes. It also extends Gillier et al.'s (2015) finding on technology trust in project teams by demonstrating that trust mediates rather than merely moderates the technology-to-performance relationship.

The moderation finding contributes to the growing literature on the contingency of technology effectiveness in project settings (Dikert et al., 2016; Serrador & Pinto, 2015). By demonstrating that agile maturity moderates the strength of the AI monitoring-to-performance indirect effect, we show that AI monitoring tools are not uniformly beneficial across agile teams their performance benefits accrue disproportionately to teams that have developed the ceremonial infrastructure, collaborative norms, and practice discipline needed to translate AI insights into performance-improving action. This finding has important implications for how organisations invest in and implement AI monitoring tools, suggesting that technology investment without corresponding investment in agile capability development is likely to produce disappointing performance returns.

- *Implications for Project Management Practice*

The findings offer several actionable implications for project managers, Scrum masters, and IT organisations. First, AI-enabled monitoring tools should be positioned not merely as project oversight mechanisms but as team performance resources — tools that, when trusted and actively used by team members, produce genuine improvements in sprint delivery quality and team collaboration. Project managers should actively involve team members in the selection, configuration, and interpretation of AI monitoring tools, building the shared ownership and transparency that are prerequisites for team trust. Second, the mediation finding underscores the importance of building team trust in AI systems as a deliberate management activity, not a passive by-product of tool deployment. This means being transparent about how AI monitoring data is collected, processed, and used in performance evaluations and actively demonstrating that AI monitoring supports rather than surveils team members. Third, the moderation finding suggests that organisations with low agile maturity should invest in agile capability development before deploying AI monitoring tools, or should implement AI monitoring tools gradually alongside agile coaching and maturity improvement programmes.

- *Limitations and Directions for Future Research*

Several limitations of this study suggest directions for future project management research. First, the cross-sectional design limits causal inference. Longitudinal research designs tracking the same agile teams across multiple sprint cycles would enable examination of how team trust in AI monitoring evolves over time and how this evolution shapes the long-term trajectory of team performance. Second, while the multi-country design enhances generalisability, it also introduces cultural heterogeneity that the study does not fully theorise or model. Future research should examine how national and organisational culture conditions the relationship between AI monitoring and team trust a question with significant practical implications for global software development organisations. Third, future studies could examine the darker side of AI monitoring in agile teams the potential for AI surveillance to undermine psychological safety, increase performance anxiety, and reduce the intrinsic motivation that is essential to agile team creativity and innovation. Fourth, the study focused on AI monitoring tools broadly defined; future research should disaggregate different types of AI monitoring tools predictive analytics, automated retrospective facilitators, intelligent burndown trackers to examine which specific capabilities drive the performance benefits identified here.

VIII. CONCLUSION

Agile software development teams are increasingly equipped with AI-enabled monitoring tools that promise real-time performance visibility, predictive risk detection, and data-driven sprint management. Whether these tools actually improve agile team performance and under what conditions has been a theoretically unresolved and empirically underexplored question in the project management literature. This study has provided a theoretically grounded and empirically rigorous answer: AI-enabled project monitoring does improve agile team performance, but it does so through the relational mechanism of team trust in AI monitoring, and its effects are significantly amplified in teams with higher levels of agile maturity. These findings extend the TOE framework to an agile team performance context, identify team trust as the critical relational pathway through which AI monitoring tools produce their team-level effects, and establish agile maturity as a boundary condition that project managers and organisations must attend to when investing in AI-enabled project delivery capabilities. As AI monitoring tools become an increasingly standard feature of agile project environments, understanding the human and organisational conditions that determine their effectiveness will become one of the defining practical and theoretical challenges of project management in the digital age.

REFERENCES

[1]. Baker, J. (2012). The technology–organization–environment framework. In Y. K. Dwivedi, M. R. Wade, & S. L. Schneberger (Eds.), Information systems theory (pp. 231–245). Springer.s

[2]. Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), Multilevel theory, research, and methods in organizations (pp. 349–381). Jossey-Bass.

[3]. Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. Journal of Management, 23(3), 239–290.

[4]. Conforto, E. C., Salum, F., Amaral, D. C., da Silva, S. L., & de Almeida, L. F. M. (2016). Can agile project management be adopted by industries other than software development? Project Management Journal, 45(3), 21–34.

[5]. Digital.ai. (2022). 16th Annual State of Agile Report. Digital.ai. https://digital.ai/resource-center/analyst-reports/state-of-agile-report/

[6]. Dikert, K., Paasivaara, M., & Lassenius, C. (2016). Challenges and success factors for large-scale agile transformations: A systematic literature review. Journal of Systems and Software, 119, 87–108.

[7]. Dybå, T., & Dingsøyr, T. (2008). Empirical studies of agile software development: A systematic

review. Information and Software Technology, 50(9–10), 833–859.
[8]. Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50.
[9]. Gillier, T., Hooge, S., & Piat, G. (2015). Framing value management for creative projects: An expansive perspective. International Journal of Project Management, 33(4), 947–960.
[10]. Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report results of PLS-SEM. European Business Review, 31(1), 2–24.
[11]. 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.
[12]. Hoegl, M., & Gemuenden, H. G. (2001). Teamwork quality and the success of innovative projects: A theoretical concept and empirical evidence. Organization Science, 12(4), 435–449.
[13]. Hu, L. T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues and applications (pp. 76–99). Sage.
[14]. Kerzner, H. (2017). Project Management: A Systems Approach to Planning, Scheduling, and Controlling (12th ed.). Wiley.
[15]. Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations. In K. J. Klein & S. W. J. Kozlowski (Eds.), Multilevel theory, research, and methods in organizations (pp. 3–90). Jossey-Bass.
[16]. Leite, L., Pimentel, J., Nogueira Rocha, C., & Melegati, J. (2020). Adopting agile practices in a global software development project: Challenges and implications. Information and Software Technology, 127, 106348.
[17]. Martinsuo, M., & Hoverfält, P. (2018). Change program management: Toward a capability for managing value-oriented, integrated multi-project change in its context. International Journal of Project Management, 36(1), 134–146.
[18]. Mir, F. A., & Pinnington, A. H. (2014). Exploring the value of project management: Linking project management performance and project success. International Journal of Project Management, 32(2), 202–217.
[19]. Moe, N. B., Dingsøyr, T., & Dybå, T. (2010). A teamwork model for understanding an agile team: A case study of a Scrum project. Information and Software Technology, 52(5), 480–491.
[20]. Mohammed, S., & Dumville, B. C. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. Journal of Organizational Behavior, 22(2), 89–106.
[21]. Müller, R., & Turner, J. R. (2007). Matching the project manager's leadership style to project type. International Journal of Project Management, 25(1), 21–32.
[22]. Muthen, B. O., Muthen, L. K., & Asparouhov, T. (2017). Regression and Mediation Analysis Using Mplus. Muthen & Muthen.
[23]. Oliveira, T., & Martins, M. F. (2011). Literature review of information technology adoption models at firm level. Electronic Journal of Information Systems Evaluation, 14(1), 110–121.
[24]. Pesamaa, O., Zwikael, O., Hair, J. F., & Huemann, M. (2021). Publishing quantitative papers with rigor and transparency. International Journal of Project Management, 39(3), 217–222.
[25]. Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. Psychological Methods, 15(3), 209–233.
[26]. Serrador, P., & Pinto, J. K. (2015). Does agile work? A quantitative analysis of agile project success. International Journal of Project Management, 33(5), 1040–1051.
[27]. Šmite, D., Moe, N. B., Hildrum, J., Gonzalez-Huerta, J., & Mendez, D. (2023). Work-from-home is here to stay: Call for flexibility in post-pandemic work policies. Journal of Systems and Software, 195, 111552.
[28]. Too, E. G., & Weaver, P. (2014). The management of project management: A conceptual framework for project governance. International Journal of Project Management, 32(8), 1382–1394.
[29]. Tornatzky, L. G., & Fleischer, M. (1990). The Processes of Technological Innovation. Lexington Books.
[30]. Turner, J. R., & Müller, R. (2005). The project manager's leadership style as a success factor on projects: A literature review. Project Management Journal, 36(2), 49–61.