Getting Familiar: Epistemic Hybrids and the Reproduction of Collaboration*

Jeffrey Peo
Saïd Business School, University of Oxford
jeffrey.peo.dphil@said.oxford.edu

June 11, 2026

Click HERE for the latest version

Abstract

Team-based knowledge work requires collaboration among members with specialized expertise. While sustaining such collaborations is difficult even in conventional settings, the introduction of artificial intelligence intensifies this challenge by bringing divergent approaches to knowledge generation and evaluation into the same team. This paper examines how epistemic hybrids – individuals whose formal training spans professional and data science domains – shape the persistence of cross-epistemic collaboration over time and its consequences for project performance. Using data from 1,331 AI-enabled consulting engagements conducted by a global management consultancy, I find that hybrid-inclusive project leadership teams are more likely to collaborate again in the future and accumulate higher levels of familiarity through repeated interaction across projects. This accumulated familiarity is positively associated with project performance, suggesting that epistemic hybrids improve coordination not only through their contributions within projects but also by shaping the relational structures through which collaboration is sustained over time.

* I am thankful to Mari Sako, Matthias Holweg, Michael Smets, and Michelle Rogan for their invaluable feedback and advice as well as to conference participants at AOM and EGOS for their helpful comments.

Introduction

Collaboration among individuals in knowledge-intensive work is not a one-shot event but an intertemporal process, in which relationships are formed, evaluated, and either reproduced or dissolved over time (Baum et al., 2005; Dahlander & McFarland, 2013; Gulati, 1995; Samila et al., 2022). Collaboration is especially valuable when it brings together complementary expertise (Faraj & Sproull, 2000; Grant, 1996; Kogut & Zander, 1992), and shared experience among team members improves coordination and performance (Ching et al., 2024; Huckman & Staats, 2011; Reagans et al., 2005). These benefits suggest that such collaborations should endure, as individuals continue to gain from one another's specialized knowledge. Yet, in practice, collaboration is often fragile. Even when actors are aligned toward a common objective, working relationships do not always persist (Dahlander & McFarland, 2013; Samila et al., 2022; Uribe et al., 2020).

This fragility is especially likely to arise in contexts characterized by epistemic heterogeneity, where collaborators differ not only in what they know but in how they generate, interpret, and evaluate knowledge (Knorr Cetina, 1999). Prior research on collaboration across occupational and knowledge boundaries highlights the coordination processes that enable individuals with different forms of expertise to work together (Bechky, 2003; Boland Jr & Tenkasi, 1995; Carlile, 2004; Hardy et al., 2005; Ungureanu et al., 2020). Yet this work largely assumes that alignment across knowledge boundaries is achievable. Under conditions of epistemic heterogeneity, that assumption becomes more difficult to sustain. Collaborators may lack shared standards for judging what constitutes valid knowledge, credible evidence, or a successful contribution. Such settings reveal a fundamental tension: complementarity in expertise does not ensure compatibility in how knowledge is produced and evaluated (Dougherty & Dunne, 2012; Forsythe, 1993). As cross-epistemic work becomes more prevalent, particularly through the integration of data science and artificial intelligence into professional work, this tension becomes increasingly consequential (Pachidi et al., 2021; Van Den Broek et al., 2021; Waardenburg et al., 2022). A central question thus arises: how can collaboration be sustained when participants lack a shared evaluative basis for assessing one another's contributions?

Despite the importance of repeated collaboration in knowledge-intensive work, research on why some collaborative relationships persist while others dissolve remains relatively limited. Existing explanations largely take one of two forms. One attributes persistence to antecedent conditions, such as shared network positions, prior ties, or complementary characteristics that shape the initial match (Burt, 2002; Gulati, 1995; Lane et al., 2021; Zhelyazkov, 2018). The other emphasizes interactional dynamics within collaboration, including trust-building, reciprocity, and helpful behavior that influence partners' willingness to re-engage (Dahlander & McFarland, 2013; Samila et al., 2022). While these perspectives explain repeated collaboration when collaborators can readily assess the value of prior interaction, they are less well suited to contexts in which participants lack a shared basis for evaluating what constitutes a successful outcome.

A separate body of research highlights the role of boundary-spanning individuals in enabling collaboration across domains of expertise. Spanning both intra- and inter-organizational contexts, this literature shows how such actors facilitate information processing, translation, and the integration of specialized knowledge, thereby improving coordination and performance within a given collaboration (Levina & Vaast, 2005; Mell et al., 2022; A. W. Richter et al., 2006; Tushman & Scanlan, 1981). However, this work primarily explains how collaborations are initiated and made effective once they occur, rather than whether individual collaborators choose to work together again when prior joint work is difficult to evaluate using shared standards.

In this paper, I shift attention from the antecedent conditions and interactional behaviors that shape repeated collaboration to the functional roles through which collaboration is organized. I examine how project team composition shapes whether collaboration is reproduced over time under conditions of epistemic heterogeneity. I argue that epistemic hybrids – boundary-spanning individuals with formal training across distinct epistemic domains – increase the likelihood that collaborators work together again. Through repeated re-engagement, these collaborations accumulate into higher levels of team familiarity, with consequences for coordination effectiveness and project performance.

I evaluate these relationships in a large management consulting firm, an epistemically diverse context in which domain expertise is increasingly integrated with data science and artificial intelligence (Christensen et al., 2013; Davenport et al., 2018; Dell'Acqua et al., 2026). The setting is well suited to this analysis because collaboration among senior members is both discretionary and frequently reconfigured rather than formally imposed. The focal firm includes partners with distinct epistemic backgrounds in management and data science, as well as individuals whose training spans these domains, enabling the identification of epistemic hybrids. Detailed personnel, activity, and financial records allow me to trace patterns of collaboration among partners over time, link project team composition to subsequent collaboration, and assess the performance implications of these repeated interactions.

The results show that epistemic hybrids play a central role in structuring collaboration over time. Project teams that include hybrids exhibit higher rates of subsequent collaboration, increasing the likelihood that collaborators re-engage. These repeated collaborations accumulate into higher levels of within-project team familiarity, which in turn is positively associated with project performance, consistent with the interpretation that relational structures emerging from repeated collaboration enhance coordination effectiveness. Notably, while some non-hybrid partner profiles are also associated with increased repeated collaboration, these collaborations are more concentrated around individual partners rather than diffusing across the broader team and, importantly, are not associated with improved performance. Together, these findings suggest that epistemic hybrids do not simply increase repeated collaboration; they help convert such repetition into team-level familiarity that improves coordination and performance.

This paper makes three contributions. First, it contributes to a large literature on collaboration (Dahlander & McFarland, 2013; Gulati, 1995; Samila et al., 2022; Uzzi & Spiro, 2005) by examining the conditions under which collaborative relationships persist over time. Prior work highlights how interactional dynamics, such as trust-building and helpful behavior, shape partners' willingness to re-engage, linking persistence to the dynamics of prior interactions (Dahlander & McFarland, 2013; Samila et al., 2022). By contrast, this paper conceptualizes the reproduction of collaboration as an outcome shaped by the functional roles individuals occupy within teams. In particular, it identifies epistemic hybrids as a structural mechanism that increases the likelihood that collaboration is reproduced across projects in settings where differences in how contributions are evaluated can create frictions. In doing so, the paper also extends boundary-spanning research beyond its traditional focus on external linkage, coordination, and knowledge integration across boundaries (e.g. Mell et al., 2022; Tushman & Scanlan, 1981), showing how boundary-spanning roles may shape whether collaboration is reproduced across projects and accumulates into team familiarity over time.

Second, the paper contributes to research on team familiarity (Ching et al., 2024; Huckman et al., 2009; Huckman & Staats, 2011), which has largely treated familiarity as an exogenous feature of team composition. This perspective builds on earlier work linking shared experience to coordination outcomes (Reagans et al., 2005). By contrast, this paper conceptualizes familiarity as an outcome of organizing. By linking role structure within projects to repeated collaboration across projects, it shows how the epistemic composition of teams shapes the accumulation of team familiarity over time. This perspective highlights familiarity not simply as a byproduct of repeated interaction, but as a consequence of how collaboration is organized.

Finally, the paper extends research on coordination in knowledge-intensive work by focusing on settings characterized by epistemic heterogeneity (e.g. Dougherty & Dunne, 2012; Pachidi et al., 2021; Van Den Broek et al., 2021; Waardenburg et al., 2022). Prior work has primarily focused on coordinating interdependent tasks among actors with complementary expertise (e.g. Bechky, 2003; Carlile, 2004). This paper highlights the distinct challenge of reconciling competing approaches to knowledge production and establishing what constitutes credible evidence. By showing how epistemic hybrids help sustain collaboration across epistemic boundaries, the paper contributes to a broader understanding of how coordination is achieved in knowledge work increasingly shaped by artificial intelligence.

Theory and Hypotheses

Collaboration as an Intertemporal Process under Epistemic Heterogeneity

In many settings, collaboration reflects a choice. In fields such as the performing arts and the sciences, actors exercise discretion in selecting partners – for instance, when creative professionals assemble teams for a production or when scientists initiate joint research projects following initial encounters (Lane et al., 2021; Uzzi & Spiro, 2005). A similar pattern extends to professional service firms, where, despite the formal staffing of junior roles, senior actors retain substantial autonomy in choosing collaborators (Maister, 2007). Following an initial engagement, collaborators in these fields, and others, often have opportunities to re-engage, transforming collaboration from a one-shot decision into an intertemporal sequence of relational choices.

The decision to sustain collaboration across projects offers several well-established advantages. Repeated ties reduce uncertainty and risk by providing actors with information about partners' reliability and capabilities, making future exchanges more predictable (Gulati, 1995; Podolny, 1994). At the same time, prior collaboration enhances coordination efficiency, as individuals develop shared understandings of roles, routines, and expectations, mitigating the startup costs associated with new partnerships (Huckman et al., 2009; Huckman & Staats, 2011). These repeated collaborations also facilitate richer information exchange and more effective knowledge integration, as familiarity with collaborators' expertise and communication styles improves the transfer and application of knowledge (Carlile, 2004; Gardner et al., 2012; Grant, 1996).

Yet, despite these benefits, sustained collaboration is the exception rather than the rule, and collaborative ties frequently decay over time, even when initial exchanges are productive (Burt, 2002; Dahlander & McFarland, 2013; Samila et al., 2022). This pattern reflects a fundamental fragility: collaboration depends not only on the value of complementary expertise, but on the ability and willingness of participants to align their expectations and coordinate their contributions over time. When this alignment breaks down – due to communication difficulties, coordination challenges, or divergent interpretations of roles, contributions, their evaluation, or even the purpose and nature of the collaboration itself – collaborators may perceive diminishing returns and choose not to re-engage.

Thus, whether these within-project benefits translate into sustained collaboration across projects depends on collaborators sharing a common basis for interpreting one another's contributions and assessing the outcomes of joint work. For example, engineers and technicians, or managers and creatives, are often able to find ways to work together because, despite differences in expertise, they share or can develop a common basis for interpreting contributions and evaluating outcomes (Bechky, 2003; Kellogg et al., 2006). This basis may be absent when collaboration occurs across epistemic cultures, where participants differ not only in what they know but in how they generate and evaluate knowledge, or what constitutes 'ground truth' (Forsythe, 1993; Knorr Cetina, 1999; Lebovitz et al., 2021). In such settings, collaborators must adjudicate between competing representations of the same underlying problem, rather than rely on the integration of distinct contributions. This challenge is increasingly salient as artificial intelligence reshapes knowledge production in professional settings, requiring collaboration across previously distinct epistemic domains (Berg et al., 2023; Dougherty & Dunne, 2012).

The integration of artificial intelligence into management consulting provides a particularly illuminating context in which these dynamics become consequential. Management consultants and data scientists possess complementary forms of expertise: consultants contribute client relationships, industry experience, and structured analytical frameworks (Kubr, 2002; McKenna, 2006; Whittington, 2019), while data scientists contribute capabilities in programming, data engineering, and statistical modeling (Avnoon, 2021; Davenport & Patil, 2012; Donoho, 2017; Fayyad & Hamutcu, 2020). This complementarity expands the range of problems firms can address, making collaboration between these groups both necessary and potentially valuable (Christensen et al., 2013; Davenport et al., 2018).

At the same time, these groups rely on distinct epistemic approaches. Consultants produce knowledge through situated engagement, using structured reasoning and interpretive synthesis to generate claims whose validity rests on contextual plausibility and narrative coherence (Glückler & Armbrüster, 2003; Mors, 2010; Reihlen & Nikolova, 2010; Werr & Stjernberg, 2003). In contrast, data scientists rely on digitized data, algorithmic modeling, and statistical inference, evaluating outputs based on predictive performance and reproducibility (Alaimo & Kallinikos, 2022; Breiman, 2001; Kitchin, 2014), often privileging predictive accuracy over interpretability and domain-grounded reasoning (Burrell, 2016; Lipton, 2018). When applied to a shared problem, these approaches can produce competing representations of both the problem and its solution, leading collaborators to differ in how they interpret contributions and evaluate outcomes.

A key implication of this divergence is that successful coordination does not necessarily produce agreement about value. Even when collaborative work is completed effectively, participants may evaluate its outcomes differently depending on the criteria they apply. For instance, a data scientist may assess a model based on predictive performance and statistical robustness, while a consultant may evaluate the same output in terms of interpretability and alignment with domain knowledge. When such differences persist, the value of collaboration becomes difficult to interpret, increasing uncertainty about the returns to future engagement. Under these conditions, even potentially valuable collaborations may fail to reproduce over time.

Existing explanations for collaboration persistence largely assume that participants can evaluate prior interactions using sufficiently shared criteria to assess the value of continued engagement. Structural accounts emphasize network position and prior ties, suggesting that actors embedded in overlapping or triadic network structures are more likely to reconnect over time (Burt, 2002; Krackhardt, 1998). Match quality explanations highlight ex ante attributes – such as complementary skills, shared interests, or colocation – that shape the quality of initial collaboration and expectations for future returns (Lane et al., 2021). Interactional accounts focus on behaviors within collaboration, including trust-building, reciprocity, helpfulness and the effective management of disagreement, which influence partners' willingness to re-engage (Dahlander & McFarland, 2013; Samila et al., 2022).

A related line of research highlights the role of boundary-spanning individuals in enabling collaboration across domains of expertise. Such actors facilitate the transfer and integration of knowledge by bridging disconnected communities and translating across different perspectives (Boland Jr & Tenkasi, 1995; Levina & Vaast, 2005; Mell et al., 2022; Tushman & Scanlan, 1981). This work primarily explains how collaborations form and are made effective within a given interaction, focusing on processes that support coordination and knowledge integration. To the extent that such processes improve the quality of collaborative work across boundaries, they may also support continued collaboration over time. However, in settings characterized by epistemic heterogeneity, where collaborators rely on differing evaluative frameworks, the challenge is not only to coordinate work but to adjudicate between competing approaches and to interpret and assess joint outcomes. Under these conditions, even well-coordinated collaborations may yield ambiguous or contested evaluations of success, limiting the extent to which coordination alone supports sustained collaboration.

Addressing this challenge requires actors who can not only coordinate across domains but also interpret and adjudicate between competing evaluative frameworks**.** Epistemic hybrids – boundary-spanning individuals whose training spans distinct epistemic domains and enables them to occupy a functionally distinct role within teams – are uniquely positioned to perform this role. In contexts where collaborators rely on differing criteria for what constitutes a valid contribution – for example, when one prioritizes predictive performance while another emphasizes interpretability or domain alignment – epistemic hybrids can help establish a common basis for understanding and assessing collaborative outputs. Rather than fully reconciling these differences, I argue such hybrids can reduce uncertainty in how outcomes are interpreted and evaluated, enabling more consistent judgments about the value of their joint work. By making the value of prior collaboration more legible, epistemic hybrids should increase the likelihood that collaborators will choose to re-engage over time. Based on this argument, I propose the following hypothesis:

H1. *The presence of an epistemic hybrid on a project leadership team will be positively associated with subsequent collaboration among team members.*

The Accumulation of Familiarity as Relational Structure

Repeated collaboration reflects not only a binary outcome but also a process that accumulates over time. Through repeated interaction, collaborators build shared experience that shapes how they understand and engage with one another in subsequent work (Ching et al., 2024; Gardner et al., 2012; Huckman et al., 2009; Huckman & Staats, 2011). Over the past two decades, team familiarity – typically defined as the extent of prior shared work experience and knowledge team members have of one another – has become a widely used construct in the study of project-based collaboration (for a detailed review, see: Muskat et al., 2022). In this sense, repeated collaboration constitutes a *flow* of relational interaction, while familiarity reflects the *stock* of shared experience that results from these repeated interactions. In settings characterized by epistemic divergence, such accumulated familiarity is particularly valuable as it reduces uncertainty in how collaborators integrate knowledge production processes and establishes shared expectations about how contributions are interpreted and outcomes are judged.

Despite its centrality, most research treats team familiarity as a pre-existing attribute of teams that shapes coordination and performance, paying limited attention to the processes through which familiarity develops (e.g. Ching et al., 2024; Gardner et al., 2012). Building on the preceding argument, I conceptualize familiarity as an endogenous outcome of repeated collaboration, where team composition shapes not only how collaboration is executed within projects, but how relational structures emerge over time. By reducing uncertainty in how collaborative outcomes are interpreted and evaluated, epistemic hybrids increase the likelihood that collaborators choose to work together again (H1), thereby generating a greater flow of repeated interactions across projects, which accumulates into a larger stock of shared experience among team members.

Because epistemic hybrids act as a structuring agent, increasing the likelihood of repeated collaboration, they are also more likely to become embedded in higher-familiarity collaboration structures over time. As a result, the presence of a hybrid on a focal project should be positively associated with team familiarity, even though such familiarity reflects accumulated interactions from prior collaborations rather than being directly generated by the hybrid within the focal team. Therefore, I hypothesize that:

> **H2.** *The presence of an epistemic hybrid on a project leadership team will be positively associated with higher levels of team familiarity.*

Implications of Relational Structure for Coordination and Performance

The presence of epistemic hybrids and the intertemporal structuring of collaboration may shape an organization's ability to execute project work that spans epistemic cultures. I next consider these implications individually and jointly. First, I examine how accumulated familiarity translates into coordination effectiveness in project-based work, linking the reproduction of collaboration across projects to performance within projects.

Coordinating knowledge among specialized experts is a core function of the modern firm (Argote & Ingram, 2000; Grant, 1996; Kogut & Zander, 1992) and, at the project-team level, the ability to do so is associated with superior performance (Faraj & Sproull, 2000; Gardner et al., 2012). However, coordination can be challenging because expertise is distributed across boundaries that impede mutual understanding (Bechky, 2003; Carlile, 2004). In settings characterized by epistemic heterogeneity, these challenges extend beyond the integration of distinct knowledge to include differences in how collaborators frame problems, apply analytical approaches, and reconcile competing representations during execution, creating misalignment that undermines effective coordination.

Team familiarity provides one mechanism through which such coordination challenges may be mitigated. Through repeated interaction, collaborators develop shared experience that shapes how they approach their work, align their efforts, and coordinate tasks (Faraj & Sproull, 2000; Reagans et al., 2005). This accumulated experience improves coordination by enabling individuals to better anticipate how collaborators apply their expertise and respond to differences in approach, reducing miscommunication and facilitating more efficient execution (Huckman et al., 2009; Reagans et al., 2005). Consistent with this view, empirical research shows that teams with greater familiarity exhibit superior performance across a range of settings, including e-sports, software development, and professional services (Ching et al., 2024; Gardner et al., 2012; Huckman et al., 2009), with effects that are particularly pronounced in complex environments (Espinosa et al., 2007). In settings characterized by epistemic heterogeneity, where collaborators must reconcile differing approaches to problem solving during execution, such accumulated experience should be especially valuable, as it enables teams to more effectively align their approaches and resolve differences in how work is executed. Accordingly, higher levels of familiarity should be associated with more effective coordination and, in turn, improved project performance.

Thus:

H3. *Team familiarity will be positively associated with project performance.*

In addition to the relational effects of familiarity, epistemic hybrids may also improve coordination directly within projects. By spanning epistemic boundaries, hybrids can anticipate points of divergence in how collaborators frame problems, apply analytical approaches, and evaluate solutions, while facilitating translation across different perspectives and forms of expertise (Boland Jr & Tenkasi, 1995; Carlile, 2004; Levina & Vaast, 2005). This may enable them to reconcile competing approaches to knowledge production, reduce misunderstandings, and facilitate the integration of distinct contributions within a given interaction. To the extent that epistemic hybrids reduce such frictions, their presence may be associated with improved coordination effectiveness within a project. In project-based settings such as the one used in this study, coordination effectiveness is reflected in project performance, which captures the extent to which teams successfully align their efforts and execute work relative to expectations (Faraj & Sproull, 2000; Gardner et al., 2012). Thus:

H4a. *The presence of an epistemic hybrid on a project team will be positively associated with project performance*.

Taken together, these arguments suggest two pathways through which epistemic hybrids may influence project performance: a direct pathway, consistent with boundary spanning literature, operating within projects through improved coordination, and an indirect pathway operating across projects through the accumulation of familiarity. While epistemic hybrids may improve coordination within a focal project by bridging epistemic boundaries and reducing misunderstandings, familiarity reflects a more durable relational foundation for coordination that extends beyond any single interaction. As collaborators engage in repeated interactions, familiarity embeds shared understandings about how work is conducted and how differences in perspective are resolved, enabling more effective coordination in subsequent projects.

To the extent that epistemic hybrids increase the likelihood of repeated collaboration (H1), they contribute to the accumulation of familiarity over time (H2). This accumulated familiarity, in turn, enhances coordination effectiveness and project performance (H3). Accordingly, epistemic hybrids are expected to influence project performance in part through the familiarity associated with repeated collaboration over time. Thus, the final hypothesis is:

> **H4b**. *The positive association between epistemic hybrid presence and project performance will be partially mediated by team familiarity.*

Data and Methods

Research Setting

This study draws on data from a global management consultancy that has been in operation for several decades and, at the time of data collection, generated more than US$1.5 billion in annual revenue. Beginning around 2016, the firm began establishing in-house data science capabilities to support the development of AI consulting services. Initially these services were limited to advising clients on AI strategy, with the in-house data science teams identifying market opportunities and developing use cases for clients based on available technologies. Yet, by 2020, the firm had established machine learning-based approaches to addressing client needs, with the data science teams leveraging client and market data to train models to inform management decisions. At the time of data collection, such AI projects represented more than 10% of global revenue and the firm had elected or hired 118 partners who had obtained degrees aligned with data science, representing roughly 30% of the partnership. Of these, 43 had no formal business education.

Despite the growth in data science headcount, most AI projects were co-led by partners with traditional management training. These non-technical partners contributed industry or practice area knowledge or had long-established client relationships. They also identified client issues, contextualized problems, and proposed approaches to address them. Meanwhile, partners with data science experience helped design the technical approach to the work, ensured that solutions integrated with client processes, and managed the junior data scientists on the team. Thus, AI services were not a separate unit within the firm, and partners with traditional management and data science backgrounds routinely worked together to develop AI-enabled solutions to client needs. While this appears to provide a clear division of expert labor within the project team, discussions with senior partners prior to data collection suggested that role boundaries were more fluid in practice, with individuals often defaulting to approaches aligned with their epistemic cultures, leading to differences in how problems are framed, analyzed, and evaluated. This organizational context therefore provides a natural setting to examine how individuals trained in different epistemic cultures collaborate in the development and delivery of AI-enabled knowledge work.

Data Sample

The consulting firm gave an exceptional level of access and provided detailed project-level data from its enterprise resource planning (ERP) system for all AI projects completed between 2020 and 2023. The initial dataset consisted of 1,368 records, of which thirty-seven appeared to reflect accounting corrections or prospecting work and were removed, leaving a final dataset of 1,331 valid projects. Data fields provided include the project start date, the project value ($USD), the planned project margin, the actual project margin, and the number of hours each contributing partner logged working on the project. Also provided was time entry data for all project types – both AI projects and traditional consulting engagements – from 2016 through 2024. While this additional data did not include the financial details provided for the AI projects, it enabled the reconstruction of collaboration patterns among partners both prior to the focal AI projects and across subsequent non-AI engagements.

In addition to project data, a roster of all active partners was provided by the firm's HR department, 316 of whom contributed to AI projects and were used in this study. Included in this data was the date each joined the firm as well as the date that they were elected partner. I supplemented these internal records with publicly available data from the firm's website and the partner's LinkedIn profiles. Fields collected through open sources included the types of undergraduate and graduate degrees earned, non-degree technical certificates completed in data science or AI, and past professional experience in data science or AI-related roles.

Together, these data allow for the analysis of the evolution of collaboration patterns across projects over time, the accumulation of team familiarity, and the consequences for project performance.

Variables

Dependent Variables

Future Collaboration Rate. To capture the extent to which collaboration is reproduced beyond the focal project, I construct a measure: Future Collaboration Rate (FCR), which reflects the proportion of within-project partner dyads that subsequently collaborate again on a future project as shown in Equation 1. By scaling the number of repeated dyadic ties by the total number of possible within-project pairings, the measure captures the extent to which relationships persist beyond the focal engagement, independent of team size. While dyads may collaborate multiple times after the focal project, the primary theoretical interest is whether the inclusion of a hybrid partner influences whether partners choose to collaborate again when future opportunities arise. Accordingly, FCR is constructed from a binary indicator at the dyadic level, where each pair of partners is coded as 1 if they subsequently collaborate and 0 otherwise. At the project level, FCR is defined as the proportion of within-project dyads that collaborate again, yielding a continuous measure bounded between 0 and 1. For robustness, I also examine alternative specifications that account for the frequency of subsequent collaborations.

Equation (1): Future Collaboration Rate (FCR)

$$\mathrm{FCR_p} = \frac{1}{\binom{n_p}{2}} \sum_{i<j} \mathrm{Reunion_{ij}}$$

where n_p denotes the number of partners on project p, such that $\binom{n_p}{2}$ is the total number of within-project dyads, and $\mathrm{Reunion}_{ij}$ indicates whether partners i and j subsequently collaborate again on a future project.

Weighted Familiarity Score. To capture the accumulated stock of prior collaboration among project team members, I construct a Weighted Familiarity Score (WFS), which reflects the extent to which partners have developed shared experience through past joint work. Whereas FCR captures the forward-looking reproduction of collaboration, WFS captures the accumulated stock of prior collaborative experience that partners bring into the focal project[1].

[1] Because WFS captures prior collaboration among partners on a focal project, it is defined only for projects involving at least two partners. Projects led by a single partner do not have an observable familiarity score. This does not imply

Consistent with Huckman et al. (2009), familiarity is constructed at the dyadic level by enumerating all $\binom{n_p}{2}$ within-project partner pairs and identifying their prior collaboration history[2]. Because familiarity reflects the accumulated depth of shared experience between collaborators, dyads characterized by more frequent interaction are likely to develop stronger shared understandings, more established routines, and greater mutual awareness (Huckman et al., 2009; Reagans et al., 2005). To account for this heterogeneity, the measure incorporates an α intensity parameter that allows for dyads that have collaborated many times to be weighted more heavily than those that have worked together only sporadically, while also recognizing diminishing marginal returns as dyads develop shared routines and mutual understanding after multiple collaborations.

Similarly, because familiarity may weaken as time passes between interactions, more recent collaborations are likely to contribute more strongly to shared understanding and coordination than temporally distant ones. Consistent with research on organizational forgetting and knowledge depreciation (Argote, 2012; Argote & Miron-Spektor, 2011; Darr et al., 1995), the measure incorporates a λ decay parameter that weights recent collaborations more heavily than older interactions. Thus, familiarity between partners who collaborate in consecutive years is higher than between those who reconnect only after several years apart. Together, these adjustments produce a recency- and intensity-weighted measure of team familiarity that captures both how often partners have worked together and how recently those collaborations occurred, while retaining a [0,1] scale. Exponential functional forms are used because they allow both temporal decay and repeated interaction effects to accumulate smoothly while exhibiting diminishing marginal influence over time. The resulting familiarity score is defined in equation 2. For robustness, I also re-estimate all models using an unadjusted familiarity measure defined as the proportion of the $\binom{n_p}{2}$ possible dyadic pairings that had previously collaborated.

---

that coordination challenges are absent in such projects. A single partner may still be coordinating the contributions of a broader project team not fully observed in the partner-level data, or integrating distinct knowledge-production approaches within the role itself.

[2] As projects last for a range of durations, in some instances partners may be collaborating on more than one project at once. In order to determine project sequence for familiarity purposes, the projects start date, as recorded in the firm's ERP was used. This approach captures both situations where partners choose to collaborate on a second project while still engaged on the first as well situations where the collaboration happens at a later date, after the initial project has been completed.

Equation (2): Weighted Familiarity Score (WFS)

$$\mathrm{WFS_p} = \frac{1}{\binom{n_p}{2}} \sum_{i<j} \left[1 - \exp\left(-\alpha \sum_{k=1}^{K_{ij}(t_p)} \exp\left(-\lambda\left(t_p - \tau_{ij,k}\right)\right) \right) \right], \quad \lambda = \frac{\ln 2}{h},\ \alpha = \frac{\ln 2}{m}$$

where n_p denotes the number of partners on project p, such that $\binom{n_p}{2}$ represents the number of within-project dyads. $K_{ij}(t_p)$ denotes the number of prior collaborations between partners i and j before the time of project p, and $\tau_{ij,k}$ denotes the timing of the k-th collaboration between the pair. The parameter h represents the temporal half-life governing how quickly familiarity decays as time elapses between collaborations, while m represents the interaction half-life governing diminishing marginal returns to repeated collaboration.

Because there is no canonical rate at which relational familiarity accumulates or decays, the intensity and decay parameters are theoretically motivated rather than empirically fixed. In the baseline specification, the temporal half-life h is set to one year and the interaction half-life m to three collaborations, implying that familiarity depreciates meaningfully in the absence of continued interaction and that the marginal coordination benefits of repeated collaboration diminish after several joint projects. Because the appropriate decay and saturation rates may vary across organizational contexts, I also examine alternative parameterizations of the temporal and interaction half-life values as a robustness check.

Project Performance. To assess coordination effectiveness, I construct a dependent variable capturing project performance relative to ex ante expectations. The measure is calculated by subtracting a project's planned margin from its realized margin, providing an observable indicator of project delivery relative to plan in project-based work (Fong Boh et al., 2007; Reagans et al., 2005), as shown in equation 3. Under the firm's fixed-fee, variable-cost model, planned margins are derived during project planning, when partners estimate the hours required from the team they expect to staff on the engagement and apply each team member's standardized internal billing rate to estimate projected labor costs. Realized margins are calculated using the same internal billing rates, but with actual hours worked rather than projected hours. Positive values therefore indicate that realized margins exceeded planned margins, suggesting lower-than-anticipated labor inputs, whereas negative values indicate that realized margins fell short of planned margins, suggesting higher-than-anticipated labor inputs. Importantly, there was no indication that the firm accounted for the presence or absence of hybrid partners when developing ex ante cost projections. The measure therefore captures performance relative to a common planning logic across projects, rather than one explicitly adjusted for hybrid team composition.

Equation (3) Project Performance

$$\text{Project Performance} = \text{Actual Margin} - \text{Planned Margin}$$

Deviations from planned margins may arise from inaccuracies in ex ante estimates, challenges during project execution, or both. At the planning stage, estimates may be inaccurate if partners underestimate the work required, misjudge the complexity of the client problem, or fail to anticipate the effort needed to integrate consulting and data science activities. During execution, deviations may arise when teams encounter miscommunication, rework, scope clarification, or difficulties integrating interdependent contributions across partners and sub-teams. In epistemically diverse teams, these challenges may be amplified by differences in how collaborators generate, interpret, and evaluate knowledge, increasing the likelihood that realized labor inputs diverge from initial plans.

Because the distribution of project performance is highly right-skewed and includes extreme values, I apply a quantile transformation prior to estimation to improve normality and stabilize variance while preserving the rank ordering of observations. As a robustness check, I re-estimate the models using the raw, untransformed measure as well as alternative transformations, including log, inverse hyperbolic sine, and Yeo–Johnson transformations.

Independent Variables

The focal independent variables capture the epistemic composition of project teams. While all firm partners belong to the same professional group – management consultants – they differ in how they are trained to generate, evaluate, and apply knowledge, reflecting distinct occupational and epistemic orientations (Bechky, 2011; Knorr Cetina, 1999). Consistent with prior research in consulting and scientific collaborations (Biancani et al., 2014; Haas & Park, 2010; A. Richter & Schmidt, 2006), I use partners' formal educational backgrounds as observable indicators of training in distinct knowledge production practices. I define three primary groups: (1) Traditional Management: Partners with undergraduate or graduate degrees in business, finance, economics, or accounting, or who hold an MBA. (2) Data Science: Partners with degrees in data science, artificial intelligence, computer science, mathematics, statistics, operations research, or electrical/computer engineering. (3) Epistemic Hybrid: Partners who meet the criteria for the data science group and who subsequently earned an MBA or equivalent management-oriented graduate degree, or partners who meet the criteria for the traditional management group who have since gained data science knowledge through technically-oriented online coursework and programs. These hybrid partners have received formal training in both computational and managerial knowledge-production practices, grounded, respectively, in statistical inference and client-centered problem framing. Rather than representing an intermediate category, hybrids reflect formal training across two epistemic cultures. Finally, some partners meet none of the above criteria. These individuals entered consulting after earning degrees in fields such as the natural or life sciences, non-computer-related engineering disciplines, or the social sciences and humanities. I refer to these individuals as non-focal partners because their training places them outside the two epistemic cultures central to this study. Although this category encompasses diverse educational backgrounds, they are analytically grouped together because they lack formal training in either of the two focal epistemic traditions. Accordingly, this group serves as a contrast category in supplemental analyses examining whether actors who span epistemic cultures play a distinctive role in structuring collaboration.

Although all partners at the firm are formally management consultants and are not separately classified by the firm, I use the terms *traditional partner, data science partner, hybrid or epistemic hybrid partner*, and *non-focal partner* to refer to partners in each epistemic group. Because the unit of analysis in this paper is at the project level, I translate the partner-level classifications into project level indicators and create a dummy variable coded 1 if at least one contributing partner on the project belongs to that group and 0 otherwise.

Control Variables

Several control variables are included to account for project attributes and partner-level experience that may influence coordination, collaboration, and the accumulation of familiarity over time. First, to account for project scale, I include a control for total project value (price, millions of USD). Larger projects typically span longer durations, providing greater opportunity for partners to interact, develop shared understandings, and adjust coordination routines. Because project value is highly right-skewed, it is log-transformed prior to analysis. Second, to account for coordination complexity, I include a control for the total number of partners collaborating on the project. Larger teams may face coordination challenges that are distinct from those arising from differences in epistemic culture and may also influence the likelihood of repeated collaboration and the development of familiarity. Finally, at the individual level, I control for heterogeneity in partners' experience by including measures of overall consulting experience (years at the firm at the time of the project) and cumulative hours spent on prior AI projects (log-transformed to correct for positive skew). As the analysis is conducted at the project level, these individual-level measures are averaged across the project team. Detailed definitions of all variables are available in the appendix (see: Appendix A, Table A1).

Empirical strategy

I estimate a series of models that examine (1) the persistence of collaboration beyond the focal project, (2) the accumulation of relational structure in the form of familiarity, and (3) the consequences of these structures for coordination effectiveness, as reflected in project performance. Because the theoretical mechanism operates through repeated collaboration and the accumulation of relational experience over time, the empirical analysis focuses on interactions among project leaders (partners), who exercise agency and have discretion in selecting collaborators and play a central role in structuring and coordinating project work. Accordingly, the measures of collaboration persistence and familiarity are constructed at the partner level and aggregated to the project level.

First, to assess how epistemic hybrids shape both the reproduction of collaboration (H1) and the accumulation of familiarity (H2), I estimate fractional logit models (Papke & Wooldridge, 1996), in which the dependent variables are the Future Collaboration Rate (FCR) and the Weighted Familiarity Score (WFS). Both the FCR and WFS variables are bounded in the [0,1] interval, making the fractional logit specification appropriate and allowing for observations at the boundaries. For robustness, I re-estimate both models using ordinary least squares (OLS) with heteroscedasticity-robust (HC3) standard errors.

In the case of familiarity, the theoretical argument links hybrid presence to the accumulation of familiarity through repeated collaboration over time. Empirically, however, hybrid presence is measured at the focal project level, whereas familiarity reflects the accumulation of prior interactions among team members. As a result, some hybrid-containing projects may involve first-time collaborations that have not yet generated familiarity, while familiarity produced through prior hybrid-containing collaborations may be observed on projects without a hybrid present. Consequently, the estimated association between hybrid presence and team familiarity likely understates the broader relationship between hybrids and the accumulation of familiarity over time.

Next, to assess the consequences of accumulated familiarity on coordination effectiveness as reflected in performance (H3), I estimate OLS models with heteroscedasticity-robust (HC3) standard errors using the quantile-transformed project performances as the dependent variable. The transformation reduces the influence of skewness and extreme values. In these models, project performance serves as an observable proxy for coordination effectiveness within the team, capturing the extent to which actual labor inputs align with ex ante project expectations, consistent with prior research in project-based coordination (Faraj & Sproull, 2000; Fong Boh et al., 2007).

Finally, to assess whether epistemic hybrids are directly associated with project performance (H4a), and whether this relationship is mediated by team familiarity (H4b), I implement a mediation framework that evaluates the indirect effect of hybrid presence on project performance via familiarity. Because the analysis combines linear and fractional logit estimators, indirect effects are estimated using a nonparametric bootstrap procedure to obtain valid confidence intervals (Imai et al., 2010). Given the observational nature of the project level data, these estimates are interpreted as evidence on whether the association between hybrid presence and coordination outcomes is accounted for by differences in accumulated familiarity across teams, rather than as a definitive test of causal mediation.

All models include year fixed effects to account for temporal shocks and changes in the firm's project portfolio and AI adoption over time, as well as standard controls for partner and project characteristics.

Results

Descriptive statistics, including the mean and standard deviation for all variables, are shown in Table 1. In cases where a transformation was applied to a variable prior to analysis, the transformed variable is used. In the sample, 41% of all projects include at least one hybrid partner.

Table 2 presents the main regression results, showing that projects that include an epistemic hybrid exhibit higher rates of subsequent collaboration among partners. Controlling for project scope and complexity, the presence of an epistemic hybrid is positively associated with future collaboration ($\beta = 0.287$, $p = 0.028$). This result remains consistent once accounting for the configuration of the remaining partners on the team ($\beta = 0.311$, $p = 0.024$). By contrast, neither traditional nor data science partner presence is consistently associated with higher rates of subsequent collaboration. When future collaboration is modeled using an intensity-weighted measure that captures repeated subsequent reunions among dyads, the estimated effect of hybrid presence remains positive but attenuates in both magnitude and statistical significance (see: Appendix A, Table A2). Together, these findings provide support for H1 and indicate that hybrid partners shape patterns of collaboration that persist beyond the focal project.

These patterns of repeated collaboration provide a foundation for the accumulation of familiarity among team members over time. As shown in Table 3, project teams that include an epistemic hybrid are associated with higher levels of familiarity ($\beta = 0.183$, $p = 0.043$), captured by a recency- and intensity-weighted measure of prior collaboration among team members. This association remains similar in magnitude when accounting for the presence and configuration of other partner types ($\beta = 0.175$, $p = 0.059$). Results remain substantively stable across alternative parameterizations of the weighted familiarity measure (see: Appendix A, Table A3). Together, these results provide support for H2.

Importantly, this estimate should be interpreted as a conservative proxy for the magnitude of the underlying theoretical relationship. Because hybrid presence is measured contemporaneously at the project level, whereas familiarity captures an accumulated stock of prior interactions, the coefficient reflects only the overlap between current hybrid presence and previously generated collaborative relationships. This temporal misalignment introduces attenuation: hybrid-containing projects may include newly assembled teams with little prior interaction, while familiarity generated through earlier hybrid-mediated collaborations may persist even in the absence of a hybrid on the focal project. Consequently, the estimated effect likely understates the broader role of hybrids in shaping the accumulation of familiarity over time.

By contrast, neither data science nor traditional partner presence exhibits a consistent association with weighted familiarity across specifications, providing no clear evidence that these partner types are linked to the accumulation of familiarity through prior reengagement, consistent with the findings on future collaboration rates.

Project Performance and Coordination Effectiveness

As shown in Table 4, Panel A, model 1, the presence of an epistemic hybrid is positively associated with project performance ($\beta = 0.131$, $p = 0.031$), consistent with more effective coordination between planned and realized project execution. However, this effect attenuates and falls below conventional significance thresholds ($\beta = 0.104$, $p = 0.129$) once the configuration of the remaining partners is incorporated, as shown in model 5. This suggests that part of the observed association reflects differences in the broader team composition and collaboration structures in which hybrids are embedded, rather than a purely direct effect of hybrid presence within the focal project. By comparison, models 2 and 3 show that the presence of a traditional partner is negatively associated with project performance ($\beta = -0.132$, $p = 0.038$) and data science partner presence is not significantly associated with project performance ($\beta = 0.017$, $p = 0.791$).

Because there are seven possible combinations of traditional, data science, and hybrid partners on a given project, I plot the levels of project performance for each configuration as predicted using the OLS models with HC3 robust standard errors and 95% confidence intervals from simulated draws (Figure 1). The figure shows that configurations including hybrid partners tend to exhibit higher predicted performance, although differences across configurations are modest and confidence intervals overlap. This pattern is consistent with the regression results, which suggest a positive but not uniformly strong association between hybrid presence and project performance.

Finally, because the OLS models use a quantile-transformed dependent variable to reduce the influence of extreme observations, I re-estimate the models using several alternative specifications of project performance for robustness (see: Appendix A, Tables A4 and A5). The results are consistent with the main findings.

Familiarity and Coordination Effectiveness

Team familiarity is positively associated with project performance. As shown in Table 4, Panel B, models 6 through 10, across project team configurations, weighted familiarity is positive and statistically significant ($\beta \approx 0.44$, $p < .001$), indicating that teams whose members have prior collaborative experience coordinate more effectively and providing support for H3. When familiarity is included alongside hybrid presence in the model, the coefficient on hybrid presence remains positive but declines in magnitude and is no longer statistically significant. This attenuation suggests that part of the association between hybrid presence and coordination effectiveness operates through the relational structure of collaboration among team members. Thus, H4a is only partially supported.

To examine this mechanism more directly, I estimate a mediation model in which the effect of hybrid presence on coordination effectiveness operates through team familiarity. As previously shown, the first stage indicates that hybrid presence is positively associated with higher levels of familiarity ($\beta = 0.183$, $p = 0.043$), corresponding to an average marginal increase of approximately four percentage points in the familiarity score. The second stage shows that familiarity is positively associated with project performance ($\beta = 0.434$, $p < 0.001$), while the total effect of hybrid presence is positive ($\beta = 0.131$, $p = 0.034$). The full results are provided in Table 5.

The mediation model is illustrated in Figure 2. The estimated indirect effect of hybrid presence on performance through familiarity is 0.019, with bootstrap confidence intervals indicating that the effect is statistically distinguishable from zero (95% CI [0.001, 0.041]). Together, these results provide support for H4b and suggest that the coordination advantage associated with hybrid partners is largely realized through the accumulation of familiarity, with more limited and less consistent evidence of a direct within-project effect once team composition and familiarity are taken into account.

As a robustness check, I re-estimate the a-path using OLS rather than a fractional-logit specification. The results are substantively unchanged (see: Appendix A, Table A6).

I also re-estimate the models using an unweighted familiarity measure, defined as the proportion of partner pairs with prior collaboration. Because most project teams include partners who have previously collaborated at some point, this measure exhibits limited variation across observations (see: Appendix A, Figure A1). Consistent with this distribution, hybrid presence is not significantly associated with the unweighted familiarity measure, although familiarity itself remains positively associated with project performance (see: Appendix A, Table A7).

Finally, to ensure that the results are not driven by specific assumptions about the decay and saturation parameters used in the weighted familiarity score, I re-estimate the mediation models using alternative parameterizations of the temporal and interaction half-life parameters (h and m). Results remain substantively stable across parameterizations (see: Appendix A, Table A3).

Examining Relational Structure

While these results indicate that projects including epistemic hybrids exhibit higher rates of subsequent collaboration, they do not reveal how such collaboration is structured. Consequently, increased collaboration could arise through repeated ties centered on hybrid partners themselves or through more diffuse patterns of collaboration across the team. To distinguish between these possibilities, I extend the analysis to dyad-level patterns of subsequent collaboration. Restricting the sample to projects that included an epistemic hybrid, I estimate dyad-level logistic regression models predicting whether a focal partner pairing collaborates again on a subsequent engagement. The key independent variable indicates whether the focal dyad includes a hybrid partner. Models include the same project level controls used in the primary analyses, and standard errors are clustered at the originating project level to account for within-project dependence among dyads. Analysis of 3,854 dyadic pairings across 464 hybrid-inclusive projects provides no evidence that hybrids primarily influence subsequent collaboration through ego-centric ties ($\beta = 0.085$, $p = 0.327$). Results are similar when the intensity of subsequent collaboration is considered ($\beta = 0.101$, $p = 0.176$).

I also examine whether hybrids specifically facilitate collaboration across epistemic cultures by focusing on dyads composed of traditional and data science partners. Using the same modeling approach, the sample includes 572 traditional–data science dyads across 263 projects, of which 43% were observed on projects that also included a hybrid partner. I find no evidence that the presence of a hybrid partner increases the specific likelihood that traditional and data science partners collaborate again on a subsequent project ($\beta = 0.286$, $p = 0.34$). Full results for both the ego-centric and cross-epistemic dyad analyses are reported in Table 6. Taken together with the finding that hybrids increase overall rates of subsequent collaboration, these results indicate that their influence does not operate through repeated ties centered specifically on the hybrid partner or through targeted cross-epistemic pairings. Instead, hybrids appear to foster more diffuse patterns of collaboration across the team, consistent with a shift in the broader relational structure rather than the persistence of specific dyadic ties.

Alternative Mechanism Test: Non-Focal Partners as Third Parties

The preceding analyses focus on the role of epistemic hybrids in shaping collaboration patterns and project performance. However, they do not address whether these effects require individuals who span the focal epistemic cultures, or whether similar outcomes could arise from partners positioned outside both cultures. To examine this alternative explanation, I analyze partners classified as non-focal: those whose educational backgrounds are unrelated to either data science or management and who therefore do not span the epistemic cultures at the center of this study.

I find that non-focal partners are positively associated with the future collaboration rate, although at lower magnitudes and falling just outside conventional significance thresholds ($\beta = 0.254$, $p = 0.070$). Similarly, I find that non-focal partners are also marginally associated with higher contemporaneous team familiarity ($\beta = 0.189$, $p = 0.052$). Taken together, the results are suggestive of non-focal partner influence on intertemporal collaboration similar to epistemic hybrids. However, applying the dyad-level analysis to non-focal partners reveals a distinct relational pattern. Unlike hybrids, non-focal partners establish collaboration patterns that are significantly ego-centric ($\beta = 0.438$, $p < 0.001$) with no evidence that they enhance collaboration between data science and traditional partners ($\beta = 0.145$, $p = 0.554$). Thus, while non-focal partners appear to have some influence on collaboration patterns, their influence operates primarily through their own bilateral relationships rather than through cross-epistemic or diffuse team-level collaboration structures, suggesting a different underlying mechanism at play. Importantly, mediation analysis offers limited evidence that the accumulated familiarity associated with non-focal partners is linked to superior project performance, in contrast to the accumulated familiarity associated with hybrids. This finding is consistent with earlier results showing no direct association between non-focal partner presence and performance. Complete results for the non-focal partner analyses are provided in the appendix (see: Appendix B, Tables B1 and B2).

Taken together, these analyses suggest that both hybrid and non-focal partners are associated with higher levels of future collaboration and accumulated familiarity but generate these collaborations through fundamentally different relational mechanisms with divergent consequences on coordination and performance. The influence of hybrids appears to be diffuse across the team rather than concentrated in ego-centric ties, whereas the influence of non-focal partners appears to be driven in part by repeated collaboration involving those partners themselves. Consequently, while the main analysis shows that hybrids are positively associated with project performance, no comparable performance effect emerges for non-focal partners. These findings suggest that the familiarity associated with hybrids differs from that associated with non-focal partners, both in how collaboration patterns emerge and in whether they are linked to project performance.

Discussion

Collaboration across epistemic boundaries is increasingly central to contemporary knowledge work, particularly as organizations integrate artificial intelligence and data science into established professional domains (Berg et al., 2023). Yet such collaborations are often fragile because participants differ not only in expertise, but in how they generate, interpret, and evaluate knowledge claims (Dougherty & Dunne, 2012; Forsythe, 1993; Knorr Cetina, 1999). This study examined how collaboration can nevertheless persist and repeat under such conditions.

Drawing on project-level data from a global management consultancy, I find that epistemic hybrids – individuals formally trained in both management and data science – increase the likelihood that collaborators re-engage after a focal project. Over time, these repeated interactions accumulate into greater team familiarity, which is positively associated with coordination effectiveness and project performance. The findings further suggest that the influence of hybrids operates less through strong direct effects on coordination within a focal interaction and more through shaping patterns of repeated collaboration that allow familiarity and shared relational understandings to accumulate over time.

This study makes several contributions to the literature. First, it contributes to research on collaboration persistence (Dahlander & McFarland, 2013; Samila et al., 2022), which similarly examines the continuation of collaborative relationships following prior interaction. Existing explanations largely attribute repeated collaboration to ex ante matching conditions or interactional dynamics that emerge during collaboration itself. Implicit in both perspectives is the assumption that collaborators share a sufficiently common evaluative basis to interpret whether a collaboration was successful. This study highlights how that assumption becomes unsustainable under epistemic heterogeneity, where collaborators may differ not only in expertise but in the criteria through which contributions and outcomes are assessed. In doing so, the paper shifts attention from repeated collaboration as primarily a function of network structure or interpersonal dynamics toward how team composition and role configuration stabilize the interpretation and evaluation of collaborative outcomes over time. Under such conditions, sustaining collaboration requires mechanisms that stabilize how collaborative outcomes are interpreted across repeated engagements.

The dyadic analyses further clarify how these evaluative conditions shape the structure of repeated collaboration over time. More specifically, they help distinguish whether hybrids sustain collaboration primarily through personal brokerage relationships or by enabling broader patterns of relational persistence across the team. If hybrids primarily operated through personal brokerage, one would expect repeated collaboration to become concentrated around the hybrid actor through ego-centric ties. Yet the analyses provide little evidence of this pattern. Instead, hybrids appear to facilitate more diffuse patterns of repeated collaboration across the broader team. This suggests that hybrids do not simply become indispensable intermediaries within collaboration networks. Rather, they appear to create conditions that enable broader relational persistence beyond direct dependence on the hybrid actor. In this sense, hybrids appear to shape the relational infrastructure of collaboration rather than simply occupying a central coordinating position within it.

These findings also extend research on boundary spanning by introducing an intertemporal perspective on how boundary-spanning roles influence coordination. Existing work has primarily examined how boundary spanners facilitate communication, translation, and knowledge integration within a focal collaboration (Levina & Vaast, 2005; Mell et al., 2022; Tushman & Scanlan, 1981). The findings in this study suggest that their influence may extend beyond any single interaction by increasing the likelihood that collaborators continue working together over time. Epistemic hybrids may therefore contribute not only by helping collaborators navigate epistemic differences within projects, but also by enabling the repeated engagements through which shared relational understandings and coordination routines gradually emerge without requiring full alignment in underlying epistemic commitments. More broadly, the findings suggest that the value of boundary-spanning roles may lie not only in episodic acts of translation, but also in shaping the relational conditions under which collaboration becomes more durable over time.

Viewed through this lens, the findings also reposition familiarity from a static attribute of teams to a relational structure that emerges through organizing processes. Prior research has typically treated familiarity as an exogenous condition carried into collaboration (e.g. Ching et al., 2024; Gardner et al., 2012; Huckman et al., 2009). By contrast, this study highlights how familiarity is itself produced through repeated patterns of collaboration that organizations partially shape through team composition and role configuration. In doing so, the paper shifts attention from familiarity as a background condition of coordination to familiarity as an organizational outcome that accumulates through the reproduction of collaboration over time. This perspective also has implications for practice, suggesting that organizations have agency in shaping future coordination capacity through decisions regarding team composition and role configuration.

Finally, this study contributes to a growing literature examining how artificial intelligence reshapes coordination and collaboration in knowledge-intensive organizations (e.g. Pachidi et al., 2021; Van Den Broek et al., 2021; Waardenburg et al., 2022). As firms increasingly integrate data scientists and domain experts into shared work, collaboration increasingly brings together actors from distinct epistemic cultures. Under such conditions, knowledge complementarity does not necessarily imply epistemic compatibility. In AI-enabled knowledge work, collaborators are often required to integrate alternative epistemic approaches to the same underlying analytic problem. Yet these collaborators may rely on different assumptions regarding valid evidence, appropriate methods, and problem framing.

Recent research has shown that these differences in how knowledge is generated, interpreted, and evaluated can create friction within collaboration, undermining participants' ability to assess one another's contributions and the success of joint work (Dougherty & Dunne, 2012; Pachidi et al., 2021; Waardenburg et al., 2022). The findings here suggest that one important challenge in AI-enabled organizations may therefore lie not only in integrating specialized expertise, but in sustaining collaboration when collaborators rely on competing evaluative standards. Importantly, the findings suggest that sustaining collaboration across epistemic boundaries may not require convergence on a shared knowledge production logic. Rather, durable collaboration may emerge through repeated interaction structures that stabilize coordination despite the persistence of epistemic difference. More broadly, the study highlights how the growing integration of artificial intelligence into professional work may increase the importance of organizational mechanisms and role structures that stabilize collaboration across enduring epistemic differences over time.

Limitations

This paper is not without limitations. The first limitation concerns the challenge of observing relational accumulation processes longitudinally within organizations. While future collaboration decisions can be directly observed, familiarity necessarily reflects an accumulated stock of prior interactions that unfolds over time. Empirically, this creates a temporal misalignment between contemporaneous hybrid presence and historically generated familiarity, likely attenuating the estimated relationship. This challenge highlights the difficulty of directly observing how repeated interactions accumulate into durable relational structures across projects and over time.

A related limitation concerns the measurement of coordination effectiveness. While project performance captures differences between planned and realized labor inputs under a fixed-fee project model, it may not fully capture coordination benefits that do not translate into time savings or cost efficiency. Accordingly, the measure should be interpreted as a partial but observable proxy for coordination effectiveness rather than a comprehensive measure of collaborative success.

A third limitation concerns the empirical strategy. While the quantitative design allows the identification of relational patterns associated with epistemic hybridity, it does not directly observe the day-to-day interactional processes through which hybrids stabilize collaboration across epistemic boundaries. Future ethnographic and process-based work may therefore help clarify how such dynamics unfold within ongoing project work.

The findings may also be shaped by the discretionary and recombinative nature of collaboration in the focal setting. Partners exercise substantial autonomy in choosing collaborators, projects are time bounded, and team recombination is frequent, allowing collaborators to discontinue ineffective relationships and reproduce those viewed as valuable. In more structurally constrained settings, where teams are more stable or hierarchies more strongly determine interaction patterns, epistemic hybrids may play a different role in sustaining collaboration. Relatedly, the setting centers on collaboration between data scientists and management professionals in AI-enabled consulting work. Although such collaborations are increasingly common in knowledge-intensive organizations, the findings do not imply that all forms of boundary spanning or expertise heterogeneity generate similar patterns of relational persistence. Future research should examine whether the mechanisms identified here extend to other forms of epistemic diversity and organizational contexts.

Finally, this study considers how epistemic hybrids shape the reproduction of collaboration and the accumulation of familiarity, but it does not examine the long-term sustainability of such roles for the individuals who occupy them. Prior research suggests that hybridity and dual identification can generate tension, role conflict, and identity strain (e.g. Ashforth & Reingen, 2014; Battilana et al., 2017; Kane & Levina, 2017). The findings here suggest that epistemic hybrids may play an important role in stabilizing collaboration across epistemic boundaries, but they also raise questions regarding the organizational and individual consequences of relying on such actors over time. Future research may therefore examine not only how epistemic hybrids facilitate coordination, but whether such roles remain sustainable as organizations become increasingly dependent on collaboration across epistemic cultures.

Conclusion

This study examined how collaboration persists under conditions of epistemic heterogeneity. Drawing on data from a global management consultancy, the findings show that epistemic hybrids increase the likelihood that collaborators re-engage over time, allowing familiarity to accumulate into a relational foundation for coordination and performance. Importantly, the findings suggest that the value of boundary-spanning roles may lie not only in resolving epistemic differences within a focal interaction but in enabling the repeated collaborations through which durable coordination structures emerge.

As organizations increasingly integrate artificial intelligence and specialized technical expertise into knowledge-intensive work, understanding how collaboration persists across epistemic boundaries will become increasingly important. By linking epistemic hybridity to the reproduction of collaboration, the accumulation of familiarity, and the coordination of knowledge work over time, this study offers a foundation for understanding how organizations sustain collaboration under conditions of epistemic diversity.

References

Alaimo, C., & Kallinikos, J. (2022). Organizations Decentered: Data Objects, Technology and Knowledge. *Organization Science*, *33*(1), 19–37. https://doi.org/10.1287/orsc.2021.1552

Argote, L. (2012). *Organizational learning: Creating, retaining and transferring knowledge*. Springer Science & Business Media.

Argote, L., & Ingram, P. (2000). Knowledge transfer: A basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes*, *82*(1), 150–169.

Argote, L., & Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. *Organization Science*, *22*(5), 1123–1137.

Ashforth, B. E., & Reingen, P. H. (2014). Functions of dysfunction: Managing the dynamics of an organizational duality in a natural food cooperative. *Administrative Science Quarterly*, *59*(3), 474–516.

Avnoon, N. (2021). Data Scientists' Identity Work: Omnivorous Symbolic Boundaries in Skills Acquisition. *Work, Employment and Society*, *35*(2), 332–349. https://doi.org/10.1177/0950017020977306

Battilana, J., Besharov, M., & Mitzinneck, B. (2017). On hybrids and hybrid organizing: A review and roadmap for future research. *The SAGE Handbook of Organizational Institutionalism*, *2*, 128–162.

Baum, J. A., Rowley, T. J., Shipilov, A. V., & Chuang, Y.-T. (2005). Dancing with strangers: Aspiration performance and the search for underwriting syndicate partners. *Administrative Science Quarterly*, *50*(4), 536–575.

Bechky, B. A. (2003). Sharing Meaning Across Occupational Communities: The Transformation of Understanding on a Production Floor. *Organization Science*, *14*(3), 312–330. https://doi.org/10.1287/orsc.14.3.312.15162

Bechky, B. A. (2011). Making organizational theory work: Institutions, occupations, and negotiated orders. *Organization Science*, *22*(5), 1157–1167.

Berg, J. M., Raj, M., & Seamans, R. (2023). Capturing value from artificial intelligence. *Academy of Management Discoveries*, *9*(4), 424–428.

Biancani, S., McFarland, D. A., & Dahlander, L. (2014). The semiformal organization. *Organization Science*, *25*(5), 1306–1324.

Boland Jr, R. J., & Tenkasi, R. V. (1995). Perspective making and perspective taking in communities of knowing. *Organization Science*, *6*(4), 350–372.

Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, *16*(3), 199–231.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 2053951715622512.

Burt, R. S. (2002). Bridge decay. *Social Networks*, *24*(4), 333–363.

Carlile, P. R. (2004). Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organization Science*, *15*(5), 555–568.

Ching, K., Forti, E., & Rawley, E. (2024). Competitive familiarity: Learning to coordinate by competing. *Organization Science*, *35*(4), 1409–1426.

Christensen, C. M., Wang, D., & Van Bever, D. (2013). Consulting on the cusp of disruption. *Harvard Business Review*, *91*(10), 106–114.

Dahlander, L., & McFarland, D. A. (2013). Ties that last: Tie formation and persistence in research collaborations over time. *Administrative Science Quarterly*, *58*(1), 69–110.

Darr, E. D., Argote, L., & Epple, D. (1995). The acquisition, transfer, and depreciation of knowledge in service organizations: Productivity in franchises. *Management Science*, *41*(11), 1750–1762.

Davenport, T. H., Libert, B., & Beck, M. (2018). Robo-Advisers are coming to consulting and corporate strategy. *Harvard Business Review*, *12*.

Davenport, T. H., & Patil, D. (2012). Data scientist. *Harvard Business Review*, *90*(5), 70–76.

Dell'Acqua, F., McFowland III, E., Mollick, E., Lifshitz, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2026). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality. *Organization Science*.

Donoho, D. (2017). 50 years of data science. *Journal of Computational and Graphical Statistics*, *26*(4), 745–766.

Dougherty, D., & Dunne, D. D. (2012). Digital science and knowledge boundaries in complex innovation. *Organization Science*, *23*(5), 1467–1484.

Espinosa, J. A., Slaughter, S. A., Kraut, R. E., & Herbsleb, J. D. (2007). Team knowledge and coordination in geographically distributed software development. *Journal of Management Information Systems*, *24*(1), 135–169.

Faraj, S., & Sproull, L. (2000). Coordinating expertise in software development teams. *Management Science*, *46*(12), 1554–1568.

Fayyad, U., & Hamutcu, H. (2020). Toward foundations for data science and analytics: A knowledge framework for professional standards. *Harvard Data Science Review*, *2*(2), 10–1162.

Fong Boh, W., Slaughter, S. A., & Espinosa, J. A. (2007). Learning from experience in software development: A multilevel analysis. *Management Science*, *53*(8), 1315–1331.

Forsythe, D. E. (1993). Engineering Knowledge: The Construction of Knowledge in Artificial Intelligence. *Social Studies of Science*, *23*(3), 445–477. https://doi.org/10.1177/030631293023003002

Gardner, H. K., Gino, F., & Staats, B. R. (2012). Dynamically integrating knowledge in teams: Transforming resources into performance. *Academy of Management Journal*, *55*(4), 998–1022.

Glückler, J., & Armbrüster, T. (2003). Bridging uncertainty in management consulting: The mechanisms of trust and networked reputation. *Organization Studies*, *24*(2), 269–297.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(S2), 109–122.

Gulati, R. (1995). Social structure and alliance formation patterns: A longitudinal analysis. *Administrative Science Quarterly*, 619–652.

Haas, M. R., & Park, S. (2010). To share or not to share? Professional norms, reference groups, and information withholding among life scientists. *Organization Science*, *21*(4), 873–891.

Hardy, C., Lawrence, T. B., & Grant, D. (2005). Discourse and collaboration: The role of conversations and collective identity. *Academy of Management Review*, *30*(1), 58–77.

Huckman, R. S., & Staats, B. R. (2011). Fluid tasks and fluid teams: The impact of diversity in experience and team familiarity on team performance. *Manufacturing & Service Operations Management*, *13*(3), 310–328.

Huckman, R. S., Staats, B. R., & Upton, D. M. (2009). Team familiarity, role experience, and performance: Evidence from Indian software services. *Management Science*, *55*(1), 85–100.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309.

Kane, A. A., & Levina, N. (2017). 'Am I still one of them?': Bicultural immigrant managers navigating social identity threats when spanning global boundaries. *Journal of Management Studies*, *54*(4), 540–577.

Kellogg, K. C., Orlikowski, W. J., & Yates, J. (2006). Life in the trading zone: Structuring coordination across boundaries in postbureaucratic organizations. *Organization Science*, *17*(1), 22–44.

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 2053951714528481.

Knorr Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science*, *3*(3), 383–397.

Kubr, M. (2002). *Management consulting: A guide to the profession*. International Labour Organization.

Lane, J. N., Ganguli, I., Gaule, P., Guinan, E., & Lakhani, K. R. (2021). Engineering serendipity: When does knowledge sharing lead to knowledge production? *Strategic Management Journal*, *42*(6), 1215–1244.

Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2021). Is AI ground truth really true? The dangers of training and evaluating AI tools based on experts' know-what. *MIS Quarterly*, *45*(3).

Levina, N., & Vaast, E. (2005). The emergence of boundary spanning competence in practice: Implications for implementation and use of information systems. *MIS Quarterly*, 335–363.

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, *16*(3), 31–57.

Maister, D. H. (2007). *Managing the professional service firm*. Simon and Schuster.

McKenna, C. D. (2006). *The world's newest profession: Management consulting in the twentieth century*. Cambridge University Press.

Mell, J. N., Van Knippenberg, D., Van Ginkel, W. P., & Heugens, P. P. (2022). From boundary spanning to intergroup knowledge integration: The role of boundary spanners' metaknowledge and proactivity. *Journal of Management Studies*, *59*(7), 1723–1755.

Mors, M. L. (2010). Innovation in a global consulting firm: When the problem is too much diversity. *Strategic Management Journal*, n/a-n/a. https://doi.org/10.1002/smj.837

Muskat, B., Anand, A., Contessotto, C., Tan, A. H. T., & Park, G. (2022). Team familiarity—Boon for routines, bane for innovation? A review and future research agenda. *Human Resource Management Review*, *32*(4), 100892.

Pachidi, S., Berends, H., Faraj, S., & Huysman, M. (2021). Make Way for the Algorithms: Symbolic Actions and Change in a Regime of Knowing. *Organization Science*, *32*(1), 18–41. https://doi.org/10.1287/orsc.2020.1377

Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics*, *11*(6), 619–632.

Podolny, J. M. (1994). Market uncertainty and the social character of economic exchange. *Administrative Science Quarterly*, 458–483.

Reagans, R., Argote, L., & Brooks, D. (2005). Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management Science*, *51*(6), 869–881.

Reihlen, M., & Nikolova, N. (2010). Knowledge production in consulting teams. *Scandinavian Journal of Management*, *26*(3), 279–289.

Richter, A., & Schmidt, S. L. (2006). Antecedents of the performance of management consultants. *Schmalenbach Business Review*, *58*, 365–391.

Richter, A. W., West, M. A., Van Dick, R., & Dawson, J. F. (2006). Boundary spanners' identification, intergroup contact, and effective intergroup relations. *Academy of Management Journal*, *49*(6), 1252–1269.

Samila, S., Oettl, A., & Hasan, S. (2022). Helpful behavior and the durability of collaborative ties. *Organization Science*, *33*(5), 1816–1836.

Tushman, M. L., & Scanlan, T. J. (1981). Boundary spanning individuals: Their role in information transfer and their antecedents. *Academy of Management Journal*, *24*(2), 289–305.

Ungureanu, P., Bertolotti, F., Mattarelli, E., & Bellesia, F. (2020). Collaboration and identity formation in strategic interorganizational partnerships: An exploration of swift identity processes. *Strategic Organization*, *18*(1), 171–211.

Uribe, J., Sytch, M., & Kim, Y. H. (2020). When friends become foes: Collaboration as a catalyst for conflict. *Administrative Science Quarterly*, *65*(3), 751–794.

Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology*, *111*(2), 447–504.

Van Den Broek, E., Sergeeva, A., Huysman Vrije, M., (2021). When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring. *MIS Quarterly*, *45*(3), 1557–1580. https://doi.org/10.25300/MISQ/2021/16559

Waardenburg, L., Huysman, M., & Sergeeva, A. V. (2022). In the Land of the Blind, the One-Eyed Man Is King: Knowledge Brokerage in the Age of Learning Algorithms. *Organization Science*, *33*(1), 59–82. https://doi.org/10.1287/orsc.2021.1544

Werr, A., & Stjernberg, T. (2003). Exploring management consulting firms as knowledge systems. *Organization Studies*, *24*(6), 881–908.

Whittington, R. (2019). *Opening strategy: Professional strategists and practice change, 1960 to today*. Oxford University Press.

Zhelyazkov, P. I. (2018). Interactions and interests: Collaboration outcomes, competitive concerns, and the limits to triadic closure. *Administrative Science Quarterly*, *63*(1), 210–247.

Tables and Figures

Table 1. Descriptive Statistics and Correlations

	Mean	Std. Dev.	1	2	3	4	5	6
1. Project Performance	-0.01	1.00	1.00					
2. Future Collaboration Rate	0.55	0.42	0.08*	1.00				
3. Weighted Familiarity Score	0.60	0.31	0.17*	0.42*	1.00			
4. Has Hybrid Partner	0.41	0.49	0.08*	0.09*	0.07*	1.00		
5. Has Traditional Partner	0.64	0.48	-0.05	-0.11*	-0.13*	-0.10*	1.00	
6. Has Data Science Partner	0.31	0.46	0.01	-0.05	-0.09*	0.03	0.00	1.00
7. Has Non-Focal Partner	0.55	0.50	0.03	0.13*	0.07*	0.04	-0.12*	-0.11*
8. Average Team Tenure (Years)	12.79	4.49	0.09*	0.08*	0.06	0.13*	-0.03	-0.11*
9. Average Team Cumulative AI Project Experience (Hours, Log)	5.54	1.03	0.06*	0.15*	0.20*	0.18*	-0.11*	0.11*
10. Number of Partners	2.49	1.72	0.04	-0.00	-0.07*	0.42*	0.35*	0.31*
11. Project Revenue (Million $USD, Log)	-1.62	1.15	0.08*	0.07*	-0.05	0.23*	0.23*	0.21*
			7	8	9	10	11	
7. Has Non-Focal Partner	0.55	0.50	1.00					
8. Average Team Tenure (Years)	12.79	4.49	0.20*	1.00				
9. Average Team Cumulative AI Project Experience (Hours, Log)	5.54	1.03	0.20*	0.04	1.00			
10. Number of Partners	2.49	1.72	0.35*	0.06*	0.21*	1.00		
11. Project Revenue (Million $USD, Log)	-1.62	1.15	0.21*	0.06*	0.19*	0.57*	1.00	

Note: The full project sample includes 1,331 projects. Future Collaboration Rate and Weighted Familiarity Score are dyad-based relational measures and are available for the projects for which the relevant collaboration histories can be observed (N = 919 and N = 978, respectively). Entries are Pearson correlations with pairwise deletion.
* $p < .05$

Table 2. Fractional Logit Models Predicting Future Collaboration Rate (FCR)

	(1)	(2)	(3)	(4)	(5)
Partner Type Indicators					
Has Hybrid Partner	0.287*				0.311*
	(0.131)				(0.137)
Has Traditional Partner		-0.254†			-0.095
		(0.154)			(0.164)
Has Data Science Partner			-0.110		-0.014
			(0.123)		(0.128)
Has Non-Focal Partner				0.254†	0.287†
				(0.140)	(0.151)
Controls					
Project Revenue (Log, $Mil.)	0.099†	0.105†	0.100†	0.102†	0.107†
	(0.058)	(0.058)	(0.057)	(0.057)	(0.058)
Number of Partners	-0.127**	-0.080*	-0.093**	-0.115**	-0.139**
	(0.037)	(0.036)	(0.035)	(0.035)	(0.043)
Average Partner Tenure (Years)	-0.015	-0.011	-0.011	-0.017	-0.024
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Average AI Experience (Log, Hrs.)	0.766**	0.756**	0.796**	0.789**	0.739**
	(0.103)	(0.105)	(0.102)	(0.103)	(0.106)
Exposure (Years)	0.727***	0.720***	0.724***	0.708***	0.705***
	(0.064)	(0.065)	(0.065)	(0.065)	(0.065)
Observations	919	919	919	919	919
Pseudo R^2 (McF)	0.109	0.108	0.107	0.108	0.112

Notes: Heteroscedasticity-robust standard errors (HC3) in parentheses
Year fixed effects are not included because Exposure (Years) captures the time between the focal project and the end of the observation window and is perfectly collinear with calendar year.
*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

Table 3. Fractional Logit Models Predicting Weighted Familiarity Score (WFS)

	(1)	(2)	(3)	(4)	(5)
Partner Type Indicators					
Has Hybrid Partner	0.183*				0.175†
	(0.090)				(0.093)
Has Traditional Partner		-0.173			-0.104
		(0.106)			(0.112)
Has Data Science Partner			-0.172*		-0.128
			(0.087)		(0.090)
Has Non-Focal Partner				0.189†	0.167
				(0.097)	(0.103)
Controls					
Project Revenue (Log, $Mil.)	-0.038	-0.034	-0.037	-0.037	-0.030
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
Number of Partners	-0.094**	-0.060*	-0.064*	-0.090*	-0.088**
	(0.028)	(0.027)	(0.027)	(0.027)	(0.032)
Average Partner Tenure (Years)	0.017	0.019	0.017	0.015	0.010
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Average AI Experience (Log, Hrs.)	0.331***	0.326***	0.347***	0.332***	0.306***
	(0.066)	(0.067)	(0.066)	(0.066)	(0.067)
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	978	978	978	978	978
Pseudo R^2 (McF)	0.025	0.025	0.025	0.025	0.028

Notes: Heteroscedasticity-robust standard errors (HC3) in parentheses
*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

Table 4. OLS Models Predicting Project Performance

Panel A. Direct Effects (without weighted familiarity)

	(1)	(2)	(3)	(4)	(5)
Partner Type Indicators					
Has Hybrid Partner	0.131*				0.104
	(0.061)				(0.069)
Has Traditional Partner		-0.132*			-0.102
		(0.064)			(0.073)
Has Data Science Partner			0.017		0.015
			(0.063)		(0.067)
Has Non-Focal Partner				-0.005	-0.002
				(0.061)	(0.069)
Controls					
Project Revenue (Log, $Mil.)	0.073*	0.075*	0.072*	0.072*	0.075*
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Number of Partners	-0.027	0.001	-0.013	-0.013	-0.015
	(0.020)	(0.019)	(0.019)	(0.019)	(0.026)
Average Partner Tenure (Years)	0.020**	0.020**	0.021**	0.021**	0.020**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Average AI Experience (Log, Hrs.)	0.021	0.015	0.026	0.026	0.013
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	1,331	1,331	1,331	1,331	1,331
Adjusted R^2	0.018	0.018	0.014	0.014	0.026

Panel B. Total Effect (including weighted familiarity)

	(6)	(7)	(8)	(9)	(10)
Weighted Familiarity Score	0.434***	0.433***	0.443***	0.448***	0.437***
	(0.107)	(0.107)	(0.106)	(0.106)	(0.107)
Partner Type Indicators					
Has Hybrid Partner	0.076				0.034
	(0.066)				(0.073)
Has Traditional Partner		-0.105			-0.115
		(0.073)			(0.081)
Has Data Science Partner			0.024		-0.002
			(0.062)		(0.069)
Has Non-Focal Partner				-0.069	-0.087
				(0.068)	(0.077)
Controls					
Project Revenue (Log, $Mil.)	0.095**	0.098**	0.094**	0.094**	0.098**
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Number of Partners	-0.025	-0.007	-0.018	-0.011	-0.003
	(0.022)	(0.022)	(0.021)	(0.021)	(0.027)
Average Partner Tenure (Years)	0.019*	0.019*	0.020*	0.022*	0.021*
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
Average AI Experience (Log, Hrs.)	0.124**	0.118**	0.130**	0.136**	0.121**
	(0.040)	(0.040)	(0.040)	(0.040)	(0.041)
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	978	978	978	978	978
Adjusted R^2	0.058	0.059	0.057	0.058	0.070

Notes: Heteroscedasticity-robust standard errors (HC3) in parentheses.
*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

Figure 1. Predicted performance based on team configuration
excluding projects with non-focal partner types (e.g. 'Non-Focal')

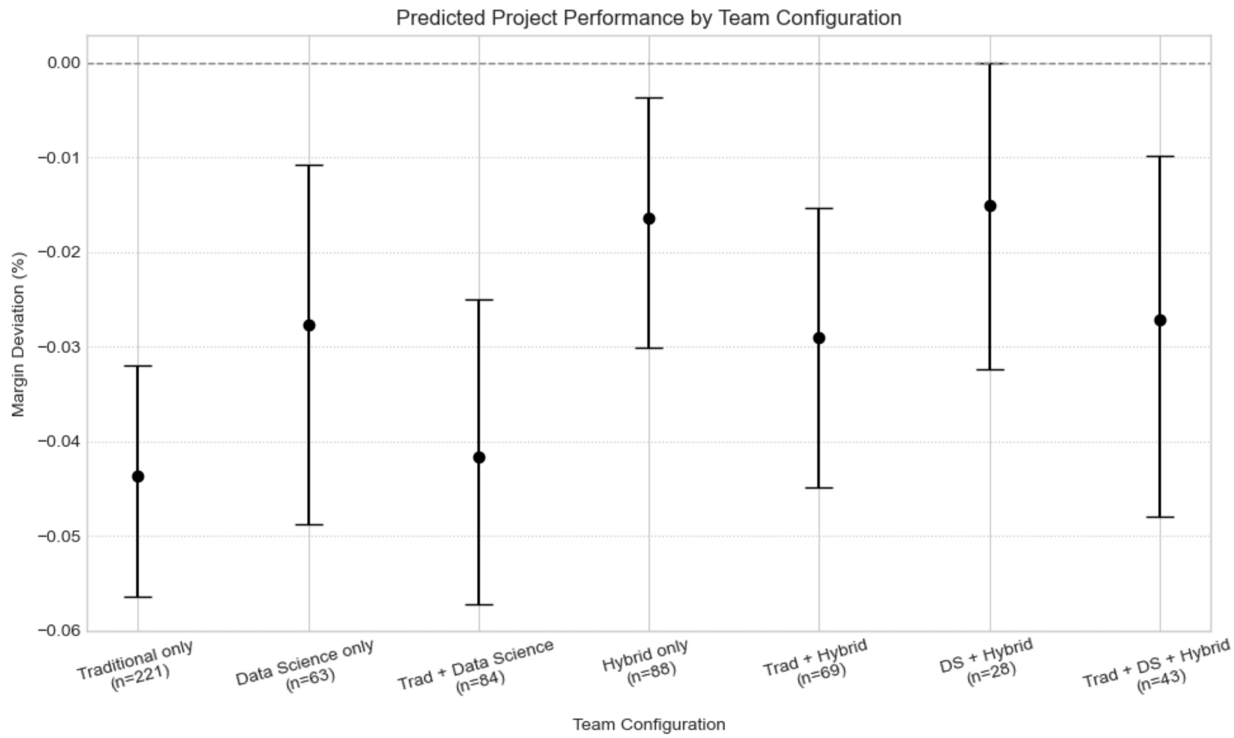


Table 5. Mediation Analysis of Hybrid Presence and Team Familiarity

Panel A. Mediation Models (Projects with Observable Familiarity: N = 978)

Model	Estimator	Focal Predictor	β	SE (HC3)	p-value
a-path	Fractional Logit	Has Hybrid → Familiarity	0.183	0.090	0.043
b-path	OLS	Familiarity → Performance	0.438	0.107	< 0.001
c'-path (Direct Effect)	OLS	Has Hybrid → Performance (with Familiarity)	0.076	0.066	0.256

Panel B. Total Effect of Hybrid Presence on Performance (N=1,331)

Model	Estimator	Sample	Focal Predictor	β	SE (HC3)	p-value	N
c-path (Total Effect)	OLS	Full Sample	Has Hybrid → Performance	0.131	0.061	0.034	1,331

Note: Controls include average partner tenure, cumulative AI experience (log hours), number of partners, and project revenue (log $ millions). All models include year fixed effects. Familiarity is defined only for projects involving two or more partners; accordingly, mediation models are estimated on the subset of projects with observable familiarity (N = 978).

Figure 2. Mediation Model Linking Hybrid Presence, Team Familiarity, and Performance.
Coefficients correspond to estimates reported in Table 5. The a-path is estimated using a fractional logit model; the b-path is estimated using OLS with HC3 robust standard errors. Direct and total effects are reported in Table 5.

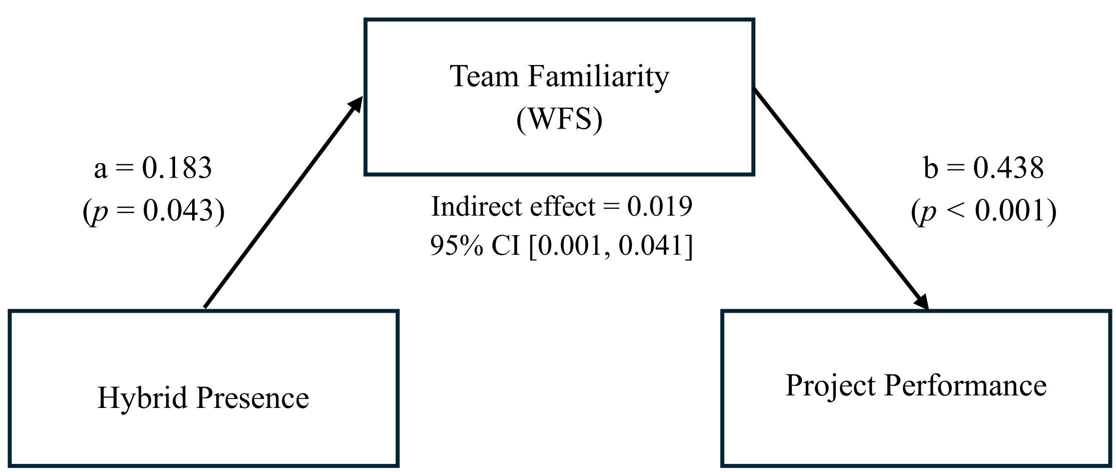


Table 6. Dyadic Collaboration Models (Hybrid Partner Effects)

	Ego-Centric Collaboration Mechanism (Hybrid Dyads)		Cross-Epistemic Collaboration Mechanism (Traditional-DS Dyads)	
	(1) Binary Reunion	(2) Weighted Reunion	(3) Binary Reunion	(4) Weighted Reunion
Dyad Includes Hybrid Partner	0.085	0.101		
	(0.086)	(0.075)		
Project Includes Hybrid Partner			0.286	0.414
			(0.301)	(0.261)
Exposure (Years)	0.343***	0.338***	0.371***	0.306***
	(0.091)	(0.079)	(0.108)	(0.087)
Other Controls	Yes	Yes	Yes	Yes
Observations (Dyads)	3,854	3,854	572	572
Observations (Projects)	464	464	263	263
Pseudo R^2	0.051	0.049	0.045	0.035

Note: Standard errors clustered at the project level in parentheses. Binary reunion indicates whether the dyad collaborates again on a subsequent project. Weighted reunion measures the intensity of subsequent collaboration using capped weighted familiarity scores. Other Controls include partner tenure, cumulative AI experience, project team size, and project revenue.

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

Appendix A. Variable Definitions and Robustness Tests

Table A1. Variable Definitions

Variable	Definition	Construction	Source
Future Collaboration Rate (FCR)	Share of within-project dyads that collaborate again on future projects	Calculated as shown in Equation (1)	Author calculation
Weighted Familiarity Score (WFS)	Recency- and intensity-weighted measure of prior collaboration among team members	Calculated using Equation (2) across all within-project dyadic partner pairs	Author calculation from collaboration history
Project Performance	Difference between realized and planned project margins	Calculated as realized margin minus planned margin, as shown in Equation (3)	ERP project financial data
Quantile-Transformed Project Performance	Rank-normalized version of Project Performance	Project Performance transformed using a rank-based quantile normalization to reduce skewness in the outcome distribution	Author calculation
Has Hybrid Partner	Indicator for whether a project includes at least one epistemic hybrid partner	Dummy variable coded 1 if ≥1 hybrid partner is present on the project team	Partner education data (LinkedIn and firm website)
Has Traditional Partner	Indicator for whether a project includes at least one traditional consulting partner	Dummy variable coded 1 if ≥1 traditional partner is present on the project team	Partner education data (LinkedIn and firm website)
Has Data Science Partner	Indicator for whether a project includes at least one partner with formal data science training	Dummy variable coded 1 if ≥1 data science partner is present on the project team	Partner education data (LinkedIn and firm website)
Has Non-Focal Partner	Indicator for whether a project includes at least one partner trained outside management and data science	Dummy variable coded 1 if ≥1 non-focal partner is present on the project team	Partner education data (LinkedIn and firm website)
Project Value	Total value of the consulting engagement (USD millions)	Log-transformed in regression models	Firm's ERP system
Team Size	Number of partners contributing to the project	Count of partners logging hours on the project	Firm's ERP system
Partner Experience	Years of consulting experience at the firm	Years between partner join date and project year; Averaged among team members	Firm's HR records
Prior AI Experience	Cumulative AI project hours logged by partners prior to focal project	Log-transformed and averaged across team members	Firm's ERP system

Robustness Tests and Alternative Specifications

This appendix reports robustness tests and alternative specifications for the future collaboration rates, main performance, familiarity, and mediation analyses. These tests examine whether the observed association between epistemic hybrid presence, familiarity, and project performance is sensitive to alternative outcome transformations, estimation methods, or parameter choices in the familiarity measure. Across specifications, the substantive conclusions remain unchanged.

Intensity-Weighted Future Collaboration Rate

The Intensity-Weighted Future Collaboration Rate (IWFCR) extends the binary FCR measure by incorporating the number of subsequent reunions among dyads. For each project p, let $\mathcal{D}_p$denote the set of unordered partner dyads on that project. For each dyad d, K_d represents the number of subsequent projects on which the same two partners collaborate again. To limit the influence of extreme values, this count is converted into an intensity weight that increases with repeated collaboration but is capped at three reunions to reflect saturation. The project-level measure is the average of these capped dyad-level weights across all dyads on the focal project:

$$\mathrm{IWFCR}_p = \frac{1}{|\mathcal{D}_{p}|}\sum_{d\in\mathcal{D}_{p}} \min\left(1, \frac{K_d}{3}\right)$$

This specification captures whether collaborative ties not only recur but also persist across multiple subsequent projects. Table A2 reports fractional logit models predicting IWFCR using the same controls included in the main analysis. The results remain substantively consistent with the FCR analysis reported in the main text.

Table A2. Fractional Logit Model Predicting Intensity-Weighted Future Collaboration Rate

	(1)	(2)	(3)	(4)	(5)
Partner Type Indicators					
Has Hybrid Partner	0.181				0.221†
	(0.131)				(0.121)
Has Traditional Partner		-0.345*			-0.161
		(0.131)			(0.142)
Has Data Science Partner			0.008		0.113
			(0.112)		(0.115)
Has Non-Focal Partner				0.350**	0.389**
				(0.122)	(0.129)
Controls					
Exposure (Years)	0.688***	0.679***	0.688***	0.666***	0.661***
	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)
Other Controls	Yes	Yes	Yes	Yes	Yes
Observations	919	919	919	919	919
Pseudo R^2 (McF)	0.098	0.100	0.096	0.101	0.104

Note: Heteroscedasticity-robust standard errors (HC3) in parentheses
Other Controls include average partner tenure, cumulative AI experience (log hours), number of partners, and project revenue (log $ millions). Year fixed effects are not included because Exposure (Years) captures the time between the focal project and the end of the observation window and is perfectly collinear with calendar year.
*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

Selection of Recency and Intensity Parameters for Weighted Familiarity Score

Next, I examine whether the mediation results are sensitive to the parameterization of the Weighted Familiarity Score (WFS). The measure incorporates two parameters governing the temporal decay of prior collaborations (h) and the saturation of repeated interactions (m). Table A3 recomputes the familiarity measure using alternative values for both parameters and re-estimates the mediation model for each specification. Across parameterizations, the estimated association between hybrid presence and familiarity remains positive and of similar magnitude, and the resulting indirect effects remain substantively comparable to the baseline specification. These results indicate that the mediation findings are not driven by the particular parameter choices used in constructing the WFS.

Table A3. Robustness of Mediation Results to Alternative WFS Parameterizations

Temporal Half-Life (h)	Interaction Half-Life (m)	Hybrid → Familiarity (β)	AME	AME 95% CI	Indirect Effect	Indirect 95% CI
1 (baseline)	3 (baseline)	0.183	0.043	(0.002, 0.084)	0.019	(0.001, 0.041)
1	2	0.182	0.040	(-0.001, 0.081)	0.017	(-0.001, 0.039)
1	4	0.179	0.043	(0.002, 0.085)	0.019	(0.001, 0.042)
2	3	0.165	0.037	(-0.004, 0.078)	0.016	(-0.002, 0.037)
3	3	0.158	0.035	(-0.006, 0.076)	0.015	(-0.003, 0.035)

Notes: Each specification recomputes the weighted familiarity score using alternative decay (h) and saturation (m) parameters before re-estimating the mediation model. Reported β values are fractional-logit coefficients for the a-path (Hybrid → Familiarity). AME values represent the average marginal effect of hybrid presence on familiarity. Indirect effects are computed as the AME multiplied by the OLS coefficient for familiarity in the outcome model. Bootstrap 95% confidence intervals are based on 5,000 resamples. N = 978 projects in all specifications.

Alternative Performance Specifications

To assess whether the relationship between hybrid presence and project performance depends on the transformation applied to the performance measure, I re-estimate the baseline models using the untransformed project performance as well as several alternative transformations designed to address skewness in the distribution. Table A4 reports OLS models using the untransformed project performance outcome, while Table A5 compares the estimated hybrid effect across multiple project performance transformations.

Table A4. OLS Models Predicting Project Performance (Untransformed)

	(1)	(2)	(3)	(4)	(5)
Partner Type Indicators					
Has Hybrid Partner	0.031†				0.027
	(0.016)				(0.019)
Has Traditional Partner		-0.017			-0.010
		(0.013)			(0.016)
Has Data Science Partner			-0.008		-0.006
			(0.015)		(0.015)
Has Non-Focal Partner				-0.000	0.000
				(0.013)	(0.014)
Controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	1,331	1,331	1,331	1,331	1,331
Adjusted R^2	0.033	0.031	0.030	0.030	0.031

Notes: Heteroscedasticity-robust standard errors (HC3) in parentheses.
*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

Table A5. Robustness of Hybrid Effect to Alternative Project Performance Transformations

Specification	**Skewness**	**Kurtosis**	**Hybrid Coef.**	**SE**	**Hybrid Effect (SE units)**
Untransformed	-5.17	51.68	0.031	0.016	0.054
Quantile Transformed	-0.00	0.71	0.131	0.061	0.034
Log-shifted	-33.79	1201.21	0.028	0.018	0.114
Inverse Hyperbolic Sine	-2.83	17.74	0.027	0.012	0.033
Yeo-Johnson	1.26	30.43	0.017	0.008	0.029

Notes: Entries report the coefficient on Has Hybrid Partner from OLS model 1 with HC3 robust standard errors. All models include the full set of controls and year fixed effects from the baseline specification. Skewness and kurtosis are reported for the dependent variable under each transformation. Because the dependent variable is transformed differently across specifications, coefficients are not directly comparable in raw units. The final column therefore reports the estimated hybrid effect expressed in standard deviation units of the transformed dependent variable. $\mathrm{N} = 1{,}331$ projects in all specifications.

Alternative Estimators

To assess whether the estimated relationship between hybrid presence and team familiarity depends on the choice of estimator, I re-estimate the a-path using an OLS specification rather than the fractional logit model used in the main analysis. As shown in Table A6, the estimated association between hybrid presence and familiarity remains positive and of similar statistical significance under both estimators.

Table A6. Alternative Estimators for the Hybrid → Familiarity Relationship

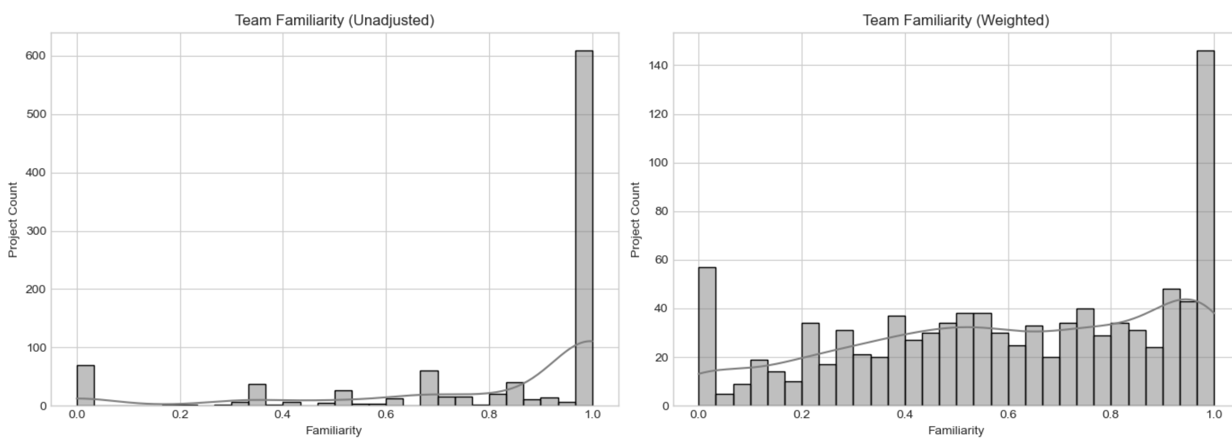
Estimator	Hybrid → Familiarity (β)	SE	p-value	N
Fractional Logit (Baseline)	0.183	0.090	0.043	978
OLS	0.043	0.021	0.045	978

Notes: All models include the same control variables and year fixed effects as the baseline specification. Standard errors are HC3 robust.

Distribution of Team Familiarity Measures

To illustrate the distributional properties of the familiarity measures used in the analysis, Figure A1 plots the distributions of both the unadjusted familiarity measure and the Weighted Familiarity Score (WFS). The unadjusted measure captures the proportion of dyads on a project with prior collaboration, while the weighted measure incorporates both the recency and frequency of past interactions between partners. As shown in the figure, the unadjusted familiarity measure is highly concentrated near the upper bound, reflecting the fact that many project teams consist largely of partners who have collaborated previously, however distantly or sparsely. In contrast, the weighted familiarity score produces a more continuous distribution across the unit interval by incorporating both the recency and frequency of prior interactions.

Figure A1. Distribution of Unadjusted and Weighted Team Familiarity Measures



Alternative Familiarity Measure

Next, I examine whether the mediation results depend on the specific construction of the familiarity measure used in the analysis. Table A7 re-estimates the mediation models using an unweighted familiarity measure defined as the proportion of dyads on a project with prior collaboration. While familiarity remains positively associated with project performance, the unweighted measure shows no relationship between hybrid presence and familiarity. As illustrated in Figure A1, the unweighted measure is highly concentrated near the upper bound and exhibits limited variance across projects, whereas the weighted familiarity score produces substantially greater dispersion by incorporating both the recency and frequency of prior interactions.

Table A7. Mediation Models Using Unweighted Familiarity

Familiarity Measure	Hybrid → Familiarity (β)	SE	p-value	Familiarity → Performance (β)	SE	p-value
Weighted Familiarity Score (baseline)	0.183	0.090	0.043	0.438	0.107	< 0.001
Unweighted Familiarity (proportion of dyads with prior collaboration)	0.019	0.145	0.894	0.331	0.123	0.007

Notes: All models include the same control variables and year fixed effects as the baseline specification. Standard errors are HC3 robust.

Appendix B. Non-Focal Partner Analyses

Dyadic Collaboration Mechanisms: Non-Focal Partners

To distinguish the role of epistemic hybrids from partners who may act as neutral brokers, I repeat the dyadic analysis focusing on non-focal partners – those whose educational backgrounds are unrelated to either management or data science. These models examine whether collaboration persistence involving non-focal partners arises primarily through ego-centric ties involving those partners or through broader cross-epistemic collaboration patterns, paralleling the approach used for hybrid partners in Table 6 of the main text. Results are reported in Table B1.

Table B1. Dyadic Collaboration Models (Non-Focal Partner Effects)

	Ego-Centric Collaboration Mechanism (Non-Focal Dyads)		**Cross-Epistemic Collaboration Mechanism (Traditional-DS Dyads)**	
	(1) Binary Reunion	(2) Weighted Reunion	(3) Binary Reunion	(4) Weighted Reunion
Dyad Includes Non-Focal Partner	0.438***	0.493***		
	(0.083)	(0.074)		
Project Includes Non-Focal Partner			0.145	0.084
			(0.245)	(0.219)
Exposure (Years)	0.456***	0.445***	0.353***	0.288***
	(0.076)	(0.065)	(0.109)	(0.090)
Other Controls	Yes	Yes	Yes	Yes
Observations (Dyads)	4,235	4,235	572	572
Observations (Projects)	699	699	263	263
Pseudo R^2	0.079	0.074	0.045	0.035

Note: Standard errors clustered at the project level in parentheses. Binary reunion indicates whether the dyad collaborates again on a subsequent project. Weighted reunion measures the intensity of subsequent collaboration using capped weighted familiarity scores. Other Controls include partner tenure, cumulative AI experience, project team size, and project revenue.
*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

Non-Focal Partners and the Familiarity-Performance Relationship

Finally, I replicate the mediation analysis presented in the main text using non-focal partner presence as the focal predictor. This test assesses whether the familiarity associated with non-focal partners is also linked to project performance. The results reported in Table B2 show that while non-focal partners are marginally associated with higher familiarity, this familiarity is not associated with improved project performance.

Table B2. Mediation Analysis of Non-Focal Presence and Team Familiarity

Model	Estimator	Focal Predictor	Coef. β	SE (HC3)	p-value	N
a-path	Fractional Logit	Has Non-Focal Partner → Familiarity	0.189	0.097	0.052	978
b-path	OLS	Familiarity → Performance	0.448	0.106	< 0.001	978
c-path (Total Effect)	OLS	Has Non-Focal Partner → Performance	0.005	0.061	0.941	1,331
c'-path (Direct Effect)	OLS	Has Non-Focal Partner → Performance (with Familiarity)	-0.069	0.068	0.307	978

Note: Controls include average partner tenure, cumulative AI experience (log hours), number of partners, and project revenue (log $ millions). All models include year fixed effects. Because the familiarity measure requires at least two partners on a project, the mediation models (a-path, b-path, and c′-path) are estimated on the subset of projects with two or more partners (N = 978). The total-effect model (c-path) uses the full sample of AI projects (N = 1,331).