Submitted to *Management Science*
manuscript

Timing Differences: Discursive Diversity and Team Performance

(Authors' names blinded for peer review)

How does cognitive diversity in a group affect its performance? Prior research suggests that cognitive diversity poses a performance tradeoff: diverse groups excel at creativity and innovation but struggle to take coordinated action. Building on the insight that group cognition is not static but is instead dynamically and interactively produced, we develop a novel conceptualization of group cognitive diversity—discursive diversity, or the degree to which the semantic meanings expressed by group members diverge from one another at a given point in time. We propose that the relationship between this time-varying measure of group cognition and team performance varies as a function of task type: discursive diversity enhances performance when groups are engaged in ideational tasks but impedes performance when they perform coordination tasks. Using the tools of computational linguistics to derive a measure of discursive diversity, and drawing on a novel longitudinal data set of intragroup electronic communications, group members' demographic traits, and performance outcomes for 117 remote software development teams on an online platform (Gigster), we find support for our theory. These results suggest that the performance tradeoff of group cognitive diversity is not inescapable: Groups can circumvent it by modulating discursive diversity to match their task requirements.

Key words: groups and teams, cognition, diversity, interaction

1. Introduction

Why do some groups perform better than others when working toward a shared goal? An extensive literature has examined this question through the lens of group diversity. The prevailing view, backed by a substantial body of empirical evidence, posits that diversity embodies a performance tradeoff: diverse groups draw on a broader set of ideas and are therefore better at discovering novel and effective solutions (e.g., Page 2008, Gibson and Vermeulen 2003), but this collective problem-solving ability comes at the expense of coordinated action, which is easier to achieve when group members' interpretations are aligned (e.g., Sørensen 2002, March 1991, Knight et al. 1999).

Scholars have uncovered this tension in a variety of contexts and at different levels of analysis. For example, cultural and ethnic diversity undermines regional and national economic growth (Alesina et al. 2003) but increases innovative capacity (Samila and Sorenson 2017) and the quality of innovation (Bernstein et al. 2019). Similarly, firms whose members hold a wide variety of cultural interpretations are better at creative innovation, while those whose members hold clashing interpretations struggle to coordinate effectively and are less profitable (Corritore et al. 2019).

The tradeoffs of convergent versus divergent thinking for group performance have been extensively studied in work on shared cognition in teams. When team members approach problems from different perspectives, they can collectively develop novel insights that no individual could have conceived of independently (Pelled et al. 1999, Amabile et al. 1996, Aggarwal and Woolley 2019). At the same time, teams can perform at a high level when each contributor understands and approaches tasks in a consistent manner, thereby enabling better communication and smoother coordination (Converse et al. 1993, Cropley 2006).

Existing research therefore suggests that teams face an inevitable tension: they can either excel at creative ideation or at coordinated execution, but not at both. We argue that this conclusion stems from the assumption that the ideas a given set of individuals brings to a group, and the behaviors these ideas catalyze, are mostly predetermined and stable over time. Yet a large body of work by interactional sociologists and social psychologists demonstrates that people produce meaning dynamically through interaction with others (Thompson and Fine 1999, Cooke et al. 2013, Berger and Luckmann 1967, Eliasoph and Lichterman 2003, Bechky 2003, Knorr Cetina and Bruegger 2002). Consistent with this view, a nascent but growing literature on the dynamics of groups suggests that groups often elude simple categorization as diverse or homogenous; rather group members often interact in ways that surface and amplify divergent ideas or instead smooth and dampen these differences over time (Cronin et al. 2011, Srikanth et al. 2016).

We build on this fundamental understanding of group cognition as dynamically and interactively produced. We argue that teams' shared cognition can fluctuate between convergence and divergence at different points in time and that these temporal shifts can influence the performance of the group as a whole. We develop this argument in two parts. First, we draw an analytical distinction between group members' private and expressed cognition, noting that people respond to interactional cues in deciding, often unselfconsciously, which of their privately held views to express in a given situation (Mobasseri et al. 2019, Goffman 1959). In group contexts, we propose that, when members discuss a given set of topics, they can express their ideas in ways that converge or diverge in meaning—independent of how similar or different their underlying, and perhaps unstated, ideas are. For example, on a product development team, members may have different understandings of what "lean" development entails, with some members focusing on minimizing waste and rework and others emphasizing the importance of failing fast and learning. We define such divergence as *discursive diversity*, or the degree to which the semantic meanings expressed by group members diverge from one another at a given point in time.

Second, we contend that the relationship between discursive diversity—a group-level cognitive construct that varies over time—and team performance varies by task type: discursive diversity boosts performance when the group is engaged in ideational tasks but undermines it when the group performs coordination tasks. We propose that conversations invoking a wide range of meanings might enable individuals to appreciate and respond to customer needs in novel ways when the team is brainstorming new product features; however, this same semantic diversity—for example, about what constitutes "lean" development—might instead lead people to talk past one another and thus fail to effectively coordinate when they are in the early stages of defining activities and negotiating roles and responsibilities.

Although there is growing interest in the role of time in research on group effectiveness (Marks et al. 2001, Volk et al. 2017, Christianson 2019), prior work—with a few exceptions (e.g., Kilduff et al. 2000)—has not explored the temporality of group cognitive alignment and its performance implications. We believe that this gap exists for methodological reasons: prior work on shared cognition has relied on static, or at best episodic, measures of group members' mental representations as reflected in self-reports. Even when implemented at multiple points in time, self-reports are ill-suited to assessing the fine-grained temporal dynamics of meaning that arise from group interaction. Thus, the emergence of shared cognition, as well as subtle shifts in cognitive diversity over time and as the team undertakes different kinds of tasks, are often obscured in studies that rely primarily on self-reports.

Using the tools of computational linguistics, we address this gap by developing a deep-learning based method for measuring the alignment, or lack thereof, of time-varying group cognition as reflected in expressed communication. We draw on longitudinal data—including intragroup electronic communications, group members' demographic traits, and performance outcomes—for 117 teams on a software development platform that matches freelance developers and project managers to projects for individual and corporate clients. Consistent with the theory we develop, our empirical analyses demonstrate that the performance benefits of divergent and convergent group cognition vary by the task the group is trying to execute. Discursive diversity reduces the likelihood of success early and late in a project milestone, when the team's tasks are more focused on coordination. In contrast, discursive diversity increases the chances of success in the middle stages of a project milestone, when the team's tasks focus more on ideation.

1.1. Group Cognitive Diversity and Performance

Team members working together toward a shared goal can diverge on a variety of dimensions such as their roles, skills, knowledge or prior experiences. An important aspect of potential divergence, which is the focus of our study, is the manner by which these individuals understand the group's objective and how they believe it should be achieved. We refer to these collective mental models as *group cognition.*[1] Given that goal-oriented teams are ultimately trying to solve a problem, group cognition can be thought of as the set of cognitive representations of the problem and how it should be solved, as well as how these representations are distributed across group members. Shared mental representations of the problem and its potential solutions represent convergent group cognition, whereas dissimilar ways of understanding the task at hand correspond to divergent group cognition. We refer to this level of divergence or convergence in group members' cognition as *group cognitive diversity.*

Considerable prior work has examined the effects of group cognitive diversity on performance. A group's joint problem-solving activity is often conceptualized as individual members searching for solutions over a stylized conceptual space. When different individuals search different areas of this space—namely, when they understand the problem differently—they are collectively more likely to find better solutions (Hong and Page 2004, Fiol and Lyles 1985, Huber 1991, Cohen and Levinthal 1990). However, this divergent search comes at the cost of increased difficulty in integrating ideas that draw on different assumptions (Converse et al. 1993). Divergent group cognition is therefore conducive to high quality problem-solving but is in tension with consistent, prompt, and coordinated execution.

Empirical evidence is generally consistent with the notion that group cognitive diversity poses a performance tradeoff.[2] Group cognitive diversity tends to boost collective creativity for two main reasons. First, when group members have divergent viewpoints, they are more likely to traverse a wider search space of ideas (Hong and Page 2004). Second, group cognitive diversity increases the probability that existing knowledge will be recombined into a novel and superior solution (Pelled et al. 1999, Amabile et al. 1996, Williams and O'Reilly 1998, de Vaan et al. 2015). Studies in

[1]Different scholars have used different terms to describe group cognition and its constituent components. (Converse et al. 1993), for example, use the term "mental model" which they define as a "knowledge structure" about the task and the ways by which team members coordinate their actions in pursuing it. Building on recent advances in research on cognition, we conceptualize individual cognition as comprised of mentally represented concepts that are held together in relationships of entailment and opposition as higher-order schematic structures (Strauss and Quinn 1997, Hannan et al. 2019). A group's cognition is convergent when team members individually activate similar schematic structures in response to the same situation.

[2]Research on the performance benefits of convergent group cognition has, however, been plagued by inconsistencies (Mohammed et al. 2010)

strategic decision-making, for example, find that decision efficiency in innovation teams was higher when members expressed frequent disagreements on innovation objectives (de Woot et al. 1977). Similarly, organizational performance was higher when executive team members expressed less consensus about strategic objectives (e.g., Bourgeois 1985). Moreover, excessive convergence in the meanings that group members convey to each other can result in various forms of groupthink, where members consensually validate each other's viewpoints at the expense of considering more accurate but contradictory information (Janis 1971, Davison and Blackman 2005).

At a same time, another body of work points to a positive relationship between aligned group cognition and effective coordination. When group members converge in the meanings they express to one another, they are more likely to find the common ground needed for coordinated action (Mohammed et al. 2000, Hinds and Bailey 2003). For example, Converse et al. (1993) found that greater overlap in team members' mental representations of group tasks and internal processes was predictive of performance for teams coordinating on complex tasks such a joint flight simulator exercise. Similarly, studies of top management teams showed that greater consensus in members' self-reported preferences about strategic firm objectives was associated with higher organizational performance (Dess 1987, Hrebiniak and Snow 1982).

1.2. Temporal Variation in Group Cognitive Diversity

Existing work sees group cognition as presenting an intractable tradeoff: groups can either innovate and learn by being cognitively divergent, or they can coordinate effectively by being cognitively convergent. In this view, maximizing creativity and innovation necessarily comes at the expense of coordination effectiveness, and vice versa. To use the imagery of individuals traversing a conceptual space, existing work generally assumes that people occupy fixed locations in this space. Yet we know that people make sense of social situations through their interactions with others (Berger and Luckmann 1967, Eliasoph and Lichterman 2003). When working together toward a shared goal, team members invariably have to take others' perspectives into account and adjust their own interpretations accordingly (Thompson and Fine 1999, Cooke et al. 2013, Knorr Cetina and Bruegger 2002). Thus, the positions group members occupy in a conceptual space—that is, the assumptions they harbor about the nature of the problem the group faces and its potential solutions—can change as they interact with one another. If group cognitive diversity is thus malleable and subject to temporal fluctuations, we propose that the performance tradeoff of group cognitive diversity is no longer inescapable.

Indeed, previous work has explored the ways by which temporal variation in group interaction relates to team performance, highlighting how temporal dynamics enable groups to oscillate between periods of ideational search and solution integration. For example, (Maznevski and Chudoba 2000) demonstrate that successful work groups fell into a rhythm that alternated between periods of intense face-to-face interaction, where the team engaged in coordination tasks, and periods of focused "solo work," where individuals focused on executing plans without much interaction. Similarly, (Bernstein et al. 2018) found that groups' performance on a complex problem-solving task improved when members exchanged information in regular but intermittent intervals, instead of constantly or not at all. The authors reasoned that the intermittent sequencing of information exchange between team members allowed individuals to alternate between ideation and coordination in a manner that benefited performance.

These studies demonstrate that the performance tradeoff of group cognitive diversity can be temporally mitigated if groups switch between different interaction modes. Whereas this previous work has focused just on the structure and temporal ordering of group interaction, we propose a different mechanism through which the tradeoff can be circumvented—through temporal changes in cognitive diversity. In particular, we posit that the performance tradeoff of group cognitive diversity can be overcome if team members can vary their levels of expressed cognitive diversity over time and in ways that match the team's task requirements. We base this argument on two important insights: first that there is a difference between what individuals subjectively experience in private and how they express their cognition in discourse; and second, that different types of tasks—specifically, ideation versus coordination tasks—require different levels of cognitive diversity for the team to perform well.

1.3. Discursive Diversity: Distinguishing Expressed from Private Cognition

Shared meaning in a group emerges interactionally, as individuals adjust their interpretations of a situation in response to the meanings expressed by others (Healey et al. 2015). Engineers and assemblers in Bechky's (2003) ethnography of a semiconductor equipment manufacturing company, for example, had to negotiate different initial understandings of technical situations, which enabled them to bridge the conceptual distances that stemmed from their different occupational experiences. In many instances such misunderstandings were only resolved when one party provided a tangible demonstration that catalyzed intense debate.

Importantly, when team members traverse cognitive distances, they do not necessarily fully align in conceptual space. Rather, they become aware of each other's different understandings and pursue interaction strategies that are mindful of and attempt to reduce conceptual distance (Hargadon and Bechky 2006). Team members thus selectively modulate which of their privately held attitudes, beliefs, and opinions they disclose to their teammates as a function of the team's social and task environment. For example, members may hold back dissenting opinions when a new domineering leader has taken over for fear of being ostracized from the group (e.g., Detert and Edmondson 2011), or they might choose to withhold novel ideas for solving a problem when a deadline is fast approaching so the team can remain focused on executing the chosen solution. Shared meaning, in other words, emerges when team members differentiate between their private and expressed cognition.

The distinction between private and expressed cognition shifts attention from what people think to how they express these thoughts in discourse. By discourse we do not simply mean the set of words expressed in language. More broadly, we use discourse to connote the underlying meanings communicated in conversation and the ways by which they reflect interaction partners' structures of knowledge and interpretation (Foucault 2002). We refer to the level of diversity in the meanings that team members convey to each other at a given point in time as *discursive diversity.*

1.4. Discursive Diversity, Task Requirements, and Team Performance

We draw on McGrath's (1991) insight that the match between group processes and the nature of the task being performed is critical for group success. Group tasks can be broadly categorized into two types: ideation tasks and coordination tasks (Bernstein et al. 2018). These task categories find broad analogues in popular task taxonomies proposed by groups and teams researchers (e.g., McGrath 1991, Marks et al. 2001, Prince and Salas 1993, Fleishman and Zaccaro 1992). For example, (McGrath 1991) proposed that team tasks can be categorized as focused on "choosing" or "executing," where "choosing" tasks involve articulating and evaluating the best options going forward, and "executing" tasks include the implementation of the chosen option and troubleshooting problems that arise in the process. Similarly, (Marks et al. 2001) proposed that teams alternate between "transition phases" and "action phases." "Transition phases" involve monitoring progress, reviewing results, and planning activities for the upcoming phases, while during "action phases," the team is focused on executing ideas and troubleshooting problems.

We propose that teams' ability to modulate their levels of discursive diversity to the task requirements they face will be predictive of team performance.[3] Specifically, we propose that discursive diversity will increase the likelihood of team success when teams are engaged in ideational tasks and will instead decrease the chances of success when teams are engaged in coordination tasks. Ideational tasks benefit from exploration of varied and unfamilar terrains in the conceptual space of ideas (Pelled et al. 1999), whereas coordination tasks require team members to be on the same page about who does what and when (Converse et al. 1993). Thus, discursive

[3]Our arguments, which are at the level of teams, have some parallels to those made by (Carnabuci and Diószegi 2015) at the individual level. They propose and find empirical support for the notion that a social network rich in structural holes boosts performance for individuals with an adaptive cognitive style, whereas a closed network is beneficial for individuals with an innovative cognitive style. Whereas they focus on the match between latent cognitive styles and the type of network in which individuals are embedded, we consider the correspondence between expressed cognitive diversity and the type of work the team is engaged in.

diversity during ideational tasks will equip group members with new ways of interpreting the shared problem and enable them to recombine ideas in ways that yield novel solutions. Conversely, discursive diversity during coordination tasks will sow confusion and make it harder for group members to find the common ground needed for smooth implementation. We therefore anticipate:

MAIN HYPOTHESIS: *Discursive diversity will increase the likelihood of success when groups are engaged in ideational tasks and will decrease the likelihood of success when groups are engaged in coordination tasks.*

1.5. Language-Based Measure of Discursive Diversity

Scholars have long speculated that team interactions and their changes over time influence the development of group cognition, but empirical investigations have lagged—in part because of limitations in available methods for exploring changes in meaning, cognition, and social interactions as they unfold over time (e.g., Fiore and Salas 2004). With a few exceptions (e.g., Kilduff et al. 2000), most prior work has conceptualized team members' cognition as relatively stable over time. Researchers have relied on surveys and interviews to assess team members' mental representations of the team's tasks and goals (e.g., Converse et al. 1993, Mohammed et al. 2000, Klimoski and Mohammed 1994), meta-knowledge about the distribution of knowledge and skills among team members (e.g., Wegner 1987), and internal team processes (e.g., Kilduff et al. 2000).

Self-reports have two key limitations. First, because they are typically administered at a single point in time or, at best, episodically, they implicitly assume that individuals' cognition is either stable or changes infrequently over the course of a team's lifespan. Consequently, the majority of studies on team cognition, whether using survey or retrospective interviews, are not designed to measure fine-grained changes in group cognition over time. Second, prior work has almost exclusively focused on self-aware and deliberative mental models as inferred from individuals' conscious reflections on team dynamics. Yet, team members interactionally produce meaning also through automated and non-reflective cognition. Indeed, what people deliberatively report is not necessarily congruent with how they unselfconsciously act (Srivastava and Banaji 2011, Healey et al. 2015).

To overcome these shortcomings and to test our main hypothesis, we develop a language-based measure of discursive diversity using the tools of natural language processing. Language reflects many important social dynamics that underlie group processes and outcomes (Lewis 2002). Generally, the similarities and differences in team members' language can reveal important information about the team's social dynamics. Interlocutors who are linguistically compatible perceive less social distance between each other than interlocutors who are linguistically divergent (Gumperz 1982, Bernstein 1971, Niederhoffer and Pennebaker 2002, Danescu-Niculescu-Mizil et al. 2012). An individual's tendency to accommodate others linguistically both affects others' evaluations (e.g., Rickford et al. 2015) and is a reflection of her self-perceived similarity with her interlocutors (e.g., Ireland et al. 2011). While these studies demonstrate the centrality of language in facilitating group interactive dynamics, they do not probe deeper into the group cognition underlying discourse.

Typically, researchers interested in how group cognition is reflected in members' linguistic exchange have focused on a single dimension of meaning contained in language, such as the concreteness of individuals' descriptions of the actions of others (Porter et al. 2016), team members' functional labels of issues (Walsh 1988), their descriptions of events as either "controllable" or "uncontrollable" (Jackson and Dutton 1988), or variation in the informational content they conveyed and in their framing of issues (Fiol 1994).

While each of these approaches highlights a potentially important dimension of meaning, they are subject to at least two critical limitations. First, each of these approaches requires that the researcher imposes her own interpretation of the meaning of the observed interaction or self-report, even though it is well-known that people's interpretations of novel information—including those of trained researchers—reflect their personal biases (e.g., Kahneman 1991, Moore et al. 2010). Thus, different researchers might interpret the same utterance from an observed interaction or self-report in different ways, such that arriving at a consistent interpretation can be challenging. Second, each of these approaches focuses only on a single dimension of meaning that team members convey to each other, privileging researchers' preconceived notions about the dimensions of meaning that are pertinent to team interaction. Focusing on a single dimension of meaning is unlikely to capture the full extent of socially relevant meaning, and thus of cognitive distance between team members, as it is reflected in their language use.

To overcome such limitations, scholars at the intersection of organization science and computational linguistics have begun to employ modern computational linguistic methods to capture more dimensions of the socially relevant meanings that group members convey in interactions with each other. These techniques can be deployed on large bodies of textual communications data that would be too complex for a human researcher to analyze. For example, Goldberg, Srivastava, and their colleagues (Goldberg et al. 2016, Srivastava et al. 2018, Doyle et al. 2017) used natural language processing techniques to develop an interactional language use model of cultural alignment based on the linguistic styles people use when communicating to their colleagues via email. They demonstrated that this language-based measure of cultural fit is predictive of consequential career outcomes such as promotion, involuntary exit, and favorable performance ratings. In a similar vein, computational analyses of the language employees use when reviewing their organizations on an online platform can be used to derive time-varying measures of cultural heterogeneity (Corritore et al. 2019). Whereas these prior studies have focused on language as a window into normative alignment at the individual level and heterogeneity in cultural perceptions at the organizational level, we instead propose to use language as a means to assessing underlying cognitive diversity at the team level.

To do so, we draw on word embedding models, a neural network-based family of unsupervised machine learning methods for representing words in a high-dimensional vector space. A word embedding model is typically trained on a large corpus of text. The specific application we use in this study relies on the continuous bag-of-words (CBOW) method wherein a two-layer neural network is trained to predict a word based on its surrounding words (Mikolov et al. 2013). Each word is then represented as a location in a shared vector space (typically comprising several hundred dimensions). The resulting dimensions of this vector space can be understood as the common latent features underlying language use in the text corpus.

Previous work demonstrates that word embedding models are particularly useful for capturing semantic relationships between words. These relationships correspond to the underlying categories of meaning that inform speakers' language use. (Garg et al. 2018), for example, demonstrate that different occupations' semantic gender associations, as inferred from word embedding models applied to English books published throughout the twentieth century, correspond to these occupations' historical gender compositions. Similarly, (Kozlowski et al. 2019) illustrate how different lifestyle activities are associated with class, race, and gender identities. Thus, word embeddings offer holistic and meaningful insights into numerous dimensions of meaning contained in language that prior methods have been unable to capture.

Let I be a team of N individuals, and W_{it} denote the set of words expressed by individual i during time period t. We define $\overline{W}_{it} = \frac{1}{|W_{it}|}\sum_w v_w$ as the embedding centroid for individual i during period t, where v_w is the embedding vector representation for word w. $\overline{W}_{it}$ represents i's embedding center of mass during time period t. This is the individual's mean position on each dimension of the embedding space as derived from her use of language during that time.

We define the embedding distance between two individuals, i and j, during time t, as the cosine distance between their respective embedding centers of mass:

$$d(W_{it}, W_{jt}) = 1 - \cos(\overline{W}_{it}, \overline{W}_{jt}) \tag{1}$$

where $\cos(A,B) = \frac{AB}{\|A\|\|B\|}$. Using this distance metric, we define the group's overall discursive diversity as the average pairwise embedding distance between all members of the group:

$$DD_t = \frac{1}{N^2}\sum_{i\in I}\sum_{j\in I} d(\overline{W}_{it}, \overline{W}_{jt}) \tag{2}$$

Our measure of discursive diversity captures the average divergence between team members' speech during a given time period. The greater this divergence, the smaller the overlap between the overall meanings expressed in each individual's language. Thus, discursive diversity reflects variation in the lenses through which individuals communicate their understanding of topics that are being discussed by their group at a given point in time. Because discursive diversity captures divergence in both the content that speakers express and the style they employ to do so, the measure captures a wide range of culturally relevant, explicit, and subtle dimensions of meaning that speakers convey to each other at a given point in time. Importantly, discursive diversity offers a direct window into team members' expressed attitudes and beliefs, as opposed to an indirect measure of latent attitudes that team members may or may not disclose to each other. Finally, our measurement approach departs from prior measures of group cognition in that it focuses on expressed, rather than conscious and self-reported, differences in cognition and in that it embraces the possibility of fine-grained temporal variation.

2. Method

2.1. Research Setting and Data

Our research setting is Gigster (gigster.com), an online platform on which freelance software developers produce on-demand software for individual and corporate clients. Unlike many two-sided platforms that match individual freelancers to clients who need help on focused, independent projects, this platform assembles individual freelance developers into temporary teams, headed by a team leader, and assigns them to longer-term projects that require complex, interdependent work. The freelancers on this platform are distributed around the globe and work on a variety of projects ranging from mobile to web application development. The projects are generally knowledge-intensive, requiring high levels of creativity, technical problem-solving, and interpersonal coordination. Software projects on this platform are significant in scope and vary in cost from tens to hundreds of thousands of dollars (and upwards of one million dollars at the extreme).

Our dataset comprises 117 teams, representing 421 unique individuals (36% female), and spans the period from early 2015 until late 2017. A typical team had 8 members and consisted of one project manager, at least one backend, frontend or "fullstack" engineer, a designer, and a user interface expert. Depending on the type of project, teams sometimes also included writers, natural language processing engineers, and other types of specialized professionals. Among teams in our data, projects lasted 159 days on average (median: 150 days) and were structured in milestone phases that lasted between one and four weeks (mean: 14 days; median: 9 days). To join the platform, professionals had to pass a variety of technical interviews designed to verify their expertise. On average, the members of an individual team represented 4.6 countries (median: 4). 42% of individuals in our sample listed their country of origin as located in North America. Another 13% hailed from Asia, followed by 12% from Europe. The remaining 23% resided in Latin America, Africa, and other parts of the world.

Because they were geographically distributed and lacked any physical office space, team members communicated almost exclusively via an online instant messaging tool called Slack. We had access to the entire set of teams' Slack archives—over 800,000 messages. Each message was timestamped and attributable (via anonymized identifiers) to its author. On average, teams exchanged 1,873 Slack messages in public channels throughout their lifespan (median: 1,220).

In addition to Slack messages, we had access to data on team member characteristics—functional role, gender, and country of origin—as well as overall team performance in meeting its various project milestones. Together, these data constitute a rich and continuous history of teams' internal dynamics and outcomes.

2.2. Dependent Variable

The timely delivery of milestones is the most critical performance measure for teams on this platform. Company executives explained that clients prioritize timeliness, and this is the key metric used to evaluate individual freelancers. The final project deliverable, as well as the deliverables for each milestone and dates for corresponding deadlines, are agreed upon between the project manager and the client before the project starts. Timely delivery signals both effective coordination among team members, as well as high output quality, since a client must approve or reject the team's deliverables at each milestone deadline. If the agreed-upon deliverables for a given milestone are deemed by the client to be of poor quality or incomplete, that milestone is marked within the company's system as delayed. Teams are allowed to proceed from one milestone phase to the next only after the client approves a given milestone's deliverables as satisfactory. Team members are paid a pre-agreed sum upon the approval of each milestone, as well as upon successful completion of the project. Members of teams that do not deliver on time may experience financial penalties or limited opportunities to join lucrative projects in the future.[4]

2.3. Independent Variables

Discursive diversity As described above, we developed a novel, interaction-based measure of the variation in the meanings that team members convey in conversation with each other. To develop this measure, we fit a word embedding model to the entire corpus of teams' Slack archives

[4]We also had access to client satisfaction scores for teams; however, after consultation with the leaders of the platform, we opted not to use this measure as a performance outcome because there is little variance in satisfaction among the clients that choose to report these scores. As we understand it, timeliness of milestone completion is the most economically consequential outcome for teams on this platform.

(vocabulary size 10,500). We pre-processed the Slack data according to standard procedures in natural language processing and trained our embeddings model using a Python implementation of Word2Vec, a popular implementation of CBOW word embeddings models. Our discursive diversity measure can be applied to time windows of varying lengths. While it provides significantly greater temporal granularity than self-report measures do, at too refined a time resolution the volume of language per individual is too sparse to generate meaningful embedding centers of mass. Because communication on Slack tends to be brief and conversational, with individual posts often comprising just a few words, we applied the measure at the daily level and then computed mean discursive diversity across project stage intervals as described below.

Given that our theory focuses on the effects of discursive diversity on different types of tasks the team performs, we turn next to describing how teams on this platform structure their work. All teams in our data operated according to the "Scrum" framework, which is a common project management approach in the software and technology industries (Schwaber 1997). Scrum allows teams to break their project work down into "mini projects" that can be completed within repeated, time-delimited iterations—so-called "sprints." Sprints typically last between two weeks and one month. Each sprint culminates in a deliverable or "milestone," which represents an incremental piece of progress toward the project's final output.

Sprints typically progress through three stages of roughly equal length, each of which involves a distinct set of tasks: sprint planning, daily Scrum work, and sprint review and integration. During sprint planning, the team creates a shared understanding of the milestone goals, the actions necessary to accomplish these goals, and the high-level tasks and responsibilities that will be assigned to team members. Overall, the planning stage focuses more on coordination than on ideation.

In the second stage of daily Scrums, the team seeks to identify how best to accomplish the goals established in the planning stage and troubleshoots new and unexpected challenges that arise as the work unfolds. Team members collectively brainstorm solutions to technical problems, ask questions, and provide feedback on each other's ideas and solutions. Although daily Scrums involve a mix of ideation and coordination, success in this stage is more about the former—that is, creative problem solving and identifying new or unanticipated ways to achieve milestone goals.

The last stage of a sprint consists of review and integration, when team members come together to integrate their individual outputs, review what has been accomplished, and discuss how to complete outstanding tasks. Overall, this last stage involves more coordination than ideation, given that teams have to close off new ideas and instead get aligned to deliver on milestone goals. To recap: in this setting, sprint planning (stage 1) and sprint review and integration (stage 3) primarily involve coordination tasks, whereas daily scrum work (stage 2) is more focused on ideation tasks.

Figure 1 provides a visual representation of the project timeline and milestone stages for a typical team.

[FIGURE 1 ABOUT HERE]

Given that our interviews of project managers and executives on the platform suggested that the three sprint stages were of roughly equal length, we used milestone-thirds as a proxy for the task that the team was engaged in at a given point in time. In other words, we assume that the first third of any milestone corresponds to sprint planning, the second third to daily Scrum work, and the final third to review and planning. We validated this assumption by qualitatively coding Slack transcripts from twenty randomly selected milestones that teams were engaged in. The tasks we coded—for example, specifying and clarifying goals, brainstorming, and monitoring progress—broadly corresponded to types of tasks that we anticipated the team would undertake across the three sprint stages.

2.4. Control Variables

Topical diversity Word embedding models capture the latent semantic features underlying language use. These latent features can include the topics being discussed—for example, when conversations revolve around backend-coding vs. design choices—but also subtle differences in the meanings expressed about these topics. Given that our arguments focus on divergence in meaning around a given set of focal topics, we sought to measure and account for the level of topical diversity present in group discussions.

We did so by training a Latent Dirichlet Allocation (LDA) topic model (Blei et al. 2003) on the entire set of teams' Slack archives. LDA topic models represent documents as distributions over topics, where the topics themselves are distributions over words. LDA "learns" the latent topics in a corpus based on the word co-occurrence patterns within documents. Treating the collection of Slack messages that an individual team member sent on a project as one document, we trained the model to identify the latent topics that team members discussed. A model with 12 topics returned what appeared to us as the most coherent and cohesive set of topics. Examples of the topics we labeled through this exercise include project management, backend engineering issues, application design choices, payments, and contract negotiations. The full set of topics, as well as illustrative key words associated with each topic, is shown in Table 1.

[TABLE 1 ABOUT HERE]

We quantified the set of topics a team member discussed during a given week as the probability distribution of her aggregated messages sent during a milestone phase over the 12 topics identified by the LDA model. The topical distance between speakers i and j during time window t can be calculated as the Hellinger distance between their respective messages' probability distributions over the latent topics for time window t. Team topical diversity during a milestone phase was measured as the mean dyadic Hellinger distance between all of a team's dyads' topic distributions during that milestone phase.

Demographic diversity We measure three types of team demographic diversity: Functional role, gender, and country of origin diversity. Each was quantified using a standardized Blau index (also referred to as an inverse Simpson or Herfindahl index), where values closer to 1 indicate greater heterogeneity and values closer to 0 indicate greater homogeneity with respect to the focal characteristic (Gibbs and Martin 1962). Our measure of discursive diversity was weakly and significantly correlated with diversity with respect to functional role ($r = 0.11, p < 0.05$) and country of origin ($r = 0.13, p < 0.01$), but not with gender diversity (see Table 4 for summary statistics and bivariate correlations). (We computed these demographic diversity measures to assess their correlation with discursive diversity. Their coefficients do not appear in our main regression models because they are subsumed in our team fixed effects.)

2.5. Analytical Strategy

We estimated linear probability models of whether or not a team achieved timely delivery of a given milestone (which we denote as "success").[5] All models described in this section include controls for topical diversity, as well as team fixed effects to account for unobserved heterogeneity in project complexity and team characteristics such as skills, ascriptive characteristics, past work experiences, and personality traits that might account for variation in team performance. The inclusion of team fixed effects also accounts for unobserved heterogeneity in team leaders' past experiences, as well as team members' past collaboration history. All models also include fixed effects for milestone phase duration to absorb unobserved differences in the likelihood of success for milestones of different lengths. We also estimate a model that includes milestone number fixed effects to account for the possibility that earlier milestones are easier to achieve than later milestones (and vice versa).

Despite the inclusion of team, milestone length, and milestone number fixed effects, which account for many potential sources of unobserved heterogeneity, we cannot fully rule out the threat of endogeneity—a point to which we return in the discussion section. We clustered errors at the team level to account for the non-independence of milestone-level observations for a given team. To test our main hypothesis, we included measures of mean discursive diversity during the first (coordination), second (ideation), and third (coordination) stages of each milestone.

[5] For robustness, we also estimated conditional logit models, which produced results that were substantively unchanged and are also reported below.

3. Results

3.1. Discursive Diversity: Validating the Word Embedding Model

As a first step, we sought to validate that our word embedding model did indeed uncover and appropriately represent the meanings expressed in team members' Slack communications. There are two common approaches to evaluating the validity of word embeddings models: most-similar queries and word analogy tasks. In most-similar queries, the model is asked to return the words that it learned to be most similar to the vector of a given target word. For example, in software development, the word "bug" usually refers to a programming issue, whereas it is more likely to refer to an insect in other contexts. Our model evaluated the most similar words to "bug" to be "issue," "crash," and "problem," demonstrating that the meaning of "bug" was accurately captured in context-relevant way. Similarly, the most similar words to "sweet" were "intense," "dope," and "yay," while the most similar words to "dude" were "man," "bro," and "yessir." These examples demonstrate an important advantage of custom-trained word embeddings over non-customized approaches to building language models: They capture not only semantic relationships but also relationships between somewhat idiosyncratic cultural schemata that are used in a given context—in this case, freelance software development teams. We conducted a wide range of most-similar queries for target words from within and outside the software development context and found that the model appeared to capture their meanings in contextually appropriate ways.

Second, (Mikolov et al. 2013) show how mathematical operations in the vector space produced by an embeddings model can be used to solve analogical reasoning problems and can serve as a further check of model validity. For example, the authors use their model to evaluate the question, "Germany is to Berlin as France is to ____?". The answer is given by $v(\text{“}Berlin'') - v(\text{“}Germany'') = x - v(\text{“}France'')$, or $x = v(\text{“}Berlin'') - v(\text{“}Germany'') + v(\text{“}France'')$. Provided that the model correctly learned vector representations of words, the answer (in this case, $x = \text{“}Paris''$) that the model returns will be given by the word vector that is closest to the coordinates that are obtained from this equation.

Applying this approach to our word embeddings model provided further evidence that our model performed well at capturing not just common semantic relationships, but also certain meanings that are idiosyncratic to the context of freelance software development. We tested our model through a number of word analogy tasks, some examples of which are shown in Table 2. In sum, both types of validity checks, most-similar queries and word analogy tasks, suggested that our model appropriately captured context-relevant semantic relationships between words.

[TABLE 2 ABOUT HERE]

To further validate our measure of discursive diversity, we qualitatively examined teams' Slack conversations during the various milestone stages. Table 3 provides sample quotations from a representative team in our data that succeeded in meeting the particular milestone from which the quotations are derived. The purpose of the table is to help illustrate how variation in discursive diversity manifests in team communication. The first page shows sample quotations from the first stage of a milestone, when discursive diversity was about one standard deviation below the mean. The team's conversation at this stage focused on sharing new information from the client and scheduling work. The second page of the table focuses on stage 2, when discursive diversity was 0.89 standard deviations above the mean. At this juncture, the team moved to troubleshooting complex technical problems, with different team members exploring and proposing different approaches. The third page of the table corresponds to the final milestone stage, when discursive diversity was 1.25 standard deviations below the mean. During this final milestone push, the team focused on fixing a few relatively minor issues and getting their interim product ready for delivery to the customer.

[TABLE 3 ABOUT HERE]

3.2. Descriptive Statistics

Table 4 shows summary statistics and bivariate correlations between the key dependent variable, milestone success, and the key independent variable, discursive diversity, as well as various control variables. Teams varied considerably with respect to average daily discursive diversity during a milestone phase ($mean: -0.04, SD: 0.83$). The average milestone phase lasted 14.02 days (SD: 18.24). The average team had 9.1 members ($SD: 4.6$). Teams were relatively diverse with respect to roles and members' countries of origin, and relatively less diverse with respect to gender.

[TABLE 4 ABOUT HERE]

Performance, measured as the timely delivery of individual milestones, was significantly and positively correlated with team role diversity ($r = 0.15, p < 0.01$) but not with topical diversity. Discursive diversity was significantly and positively correlated with team role diversity ($r = 0.11, p < 0.05$), and team diversity with respect to members' countries of origin ($r = 0.13, p < 0.01$). This makes intuitive sense, as larger teams that encompass individuals with more diverse professional or cultural backgrounds are more likely to preserve a wider array of interpretations of various topics than less demographically diverse teams. Moreover, the correlation between discursive diversity and topical diversity was weak and only marginally significant ($p < 0.1$). This suggests that discursive diversity captures variation in the meanings that team members convey in conversations with each other above and beyond the specific topics being discussed.

Finally, we examined the degree to which discursive diversity varies over time within teams. Figure 2, panel A, plots team discursive diversity as a function of team life stage, where life stages of 0 and 1 correspond to the start and end points of the project, respectively. Each of the grey lines in the figure represents one team from a sample of 20 randomly selected teams. Similarly, Figure 2, Panel B, plots team discursive diversity as a function of milestone life stage for 20 randomly selected milestone phases, where life stage 0 corresponds to the start of a milestone phase and stage 1 corresponds to the milestone due date. The blue lines represent the mean level of discursive diversity at a given team life stage (Panel A) and the mean level of discursive diversity at a given milestone life stage (Panel B), respectively.

[FIGURE 2 ABOUT HERE]

Figure 2 illustrates the core advantage of our approach to measuring cognitive diversity relative to traditional survey-based measures: Even if collected at a few points in time during a team's life cycle, self-reports of cognitive diversity would simply be unable to capture the fine-grained temporal variation that language-based measures can uncover.

Figure 3 shows the distribution of teams' mean levels of daily discursive diversity. As the plot indicates, discursive diversity varied substantially not only within, but also across, teams.

[FIGURE 3 ABOUT HERE]

Main Results

Table 5 reports our main results. First, we explore performance as a function of topical diversity, without taking discursive diversity into account. As Model 1 demonstrates, the extent to which individuals discuss a breadth of topics is unrelated to team performance. Next, we investigated whether teams' mean discursive diversity *across the milestone as a whole* is predictive of performance. Model 2 shows that, controlling for diversity in the topics that team members discuss, the relationship between mean diversity levels and performance is not significant.

[TABLE 5 ABOUT HERE]

The picture changes, however, once we take teams' relative life stage within a milestone, and, by proxy, type of task being undertaken, into account. Our main hypothesis suggests that discursive diversity will be positively associated with the chances for milestone success when it manifests in the second milestone stage (ideation), and it will reduce the likelihood of success in the first and third stages (coordination). The results in Model 3 are consistent with this expectation. This same pattern of results persists in Model 4, which includes fixed effects not only for milestone length but also for milestone number.[6] Together, these results provide strong support for our main hypothesis.

As Figure 4 illustrates, the magnitude of these effects is significant and economically meaningful. The left panel suggests that, all else equal, a team one standard deviation below the mean in

[6] In (unreported) robustness checks, we found the same basic pattern, where discursive diversity was negatively related to performance in the early and late stages, but positively related to performance in the middle stages of a milestone phase, for different splits of milestone phases—for example, when time was segmented into quarters or fifths—though results were sharper for the first and last milestone and attenuated in some cases for intervening milestones.

discursive diversity during the first stage of a milestone has a 70 percent chance of achieving milestone success, whereas for a team one standard deviation above the mean, the chances of success drop to just over 60 percent. The magnitude of the effect is comparable but in the opposite direction in the middle panel, which depicts a positive relationship between discursive diversity in the second milestone stage and team success. Finally, the magnitude and direction of the effect in the third panel, which focuses on the third stage, is comparable to that found in the first stage.

As a robustness check, we estimated conditional logit models instead of linear probability models. These results, which are substantively unchanged from those in 5, appear in Table 6.

[TABLE 6 ABOUT HERE]

4. Discussion

The goal of this article has been to bring conceptual clarity and new empirical evidence to bear on a longstanding question in organizational theory: how does cognitive diversity among members of a group influence their performance? Prior work has assumed that group cognitive diversity confers the benefits of creativity and innovation but, at the same time, imposes the costs of misalignment and strained coordination (Milliken and Martins 1996). Thus, for a given set of members, groups can shine at creativity or at implementation but generally not at both.

Building on the insight that meaning is produced collectively and dynamically through interaction (Berger and Luckmann 1967) and that interactions between group members enable them to adjust their understandings of shared problems and potential solutions (Thompson and Fine 1999), we developed a theoretical account of how groups can circumvent the performance tradeoff of group cognitive diversity. We introduced a novel, time-varying construct of group cognition, *discursive diversity*, which reflects dissimilarities in semantic meanings expressed by group members at a given point in time. Using a deep-learning linguistic method (Mikolov et al. 2013) and data from 117 software development teams on an online platform, we theorized and found empirical support for the notion that discursive diversity's relationship to performance is contingent on the nature of the group's task: it is positive when the group is engaged in ideational tasks and negative when the group performs coordination tasks.

4.1. The Role of Time in Group Processes

Findings from this paper challenge the prevailing view in research on groups and teams of diversity as a "double-edged sword" that necessarily aids creativity at the expense of coordination (Milliken and Martins 1996). By bringing in the roles of time, group interactions, and collective meaning making through discourse, we demonstrate that groups can escape the performance tradeoff of group cognitive diversity by modulating their levels of discursive diversity to match their task requirements over time.

We also make other noteworthy contributions to the burgeoning literature on the role of time in group processes (Ballard et al. 2008, Cronin et al. 2011, Srikanth et al. 2016). Prior work has importantly highlighted the different stages through which teams progress, how group cognition shifts across stages, and the implications of these changes for group learning, task conflict, and performance (Kelly and McGrath 1985, Gersick 1991, Jehn et al. 1999). Yet these models generally assume that group cognition is relatively stable within a given stage. Our findings, as highlighted in Figure 2, suggest the need to complicate these accounts: cognitive diversity, as reflected in the discourse of a group, varies considerably across teams at a given life stage. Our results also point to the limitations of self-reports, which are ill-suited to surfacing fine-grained temporal variation in group cognitive diversity, and highlight the value of digital trace data as a window into the dynamics of collective cognition.

More recent work on temporal dynamics in groups has revealed how the sequencing of interactions relates to performance. (Maznevski and Chudoba 2000), for example, find that the temporal rhythms of teams—the cadence of intense interpersonal meetings versus shorter interactions over a range of media—can importantly shape performance outcomes. In a similar vein, (Bernstein et al. 2018) demonstrate that intermittent, rather than continuous, group interactions improve average group performance while maintaining high maximum performance. Whereas these prior studies have focused on the structure and temporal ordering of group interactions, we instead draw attention to the content of group interactions and demonstrate the utility of natural language processing and machine learning methods in uncovering new facets of group cognitive diversity.

Finally, we conjecture that discursive diversity may play an important role in the emergence of other facets of group cognition. A leading candidate is group transactive memory—a repository that emerges within a team to encode, organize, and share knowledge from group members' different domains of expertise (Wegner 1987, Ren and Argote 2011, Reagans et al. 2016, Aggarwal and Woolley 2019). Transactive memory systems are more likely to emerge, and are more effective, when individuals engage in repeated interactions and form close relationships (Wegner et al. 1991). We believe that a promising path for future research is in investigating how discursive diversity influences this process. High levels of discursive diversity may, for example, enable teams to more efficiently understand and map the full range of knowledge, expertise, and skills that exist among members. Yet it may also make it harder for teams to successfully encode and retrieve this knowledge. It may even be possible to develop language-based measures of a group's transactive memory system so that its interrelationship with discursive diversity can be directly examined.

4.2. Collective Meaning Making in Groups

Next, we believe that the methodological approach we develop to measure discursive diversity can be extended beyond the context of goal-directed teams to understand collective meaning making in groups more generally—from children negotiating social roles on the playground to analysts making sense of turbulent financial markets (Knorr Cetina and Bruegger 2002). For example, we suspect that discursive diversity may be central to what (Collins 2014) refers to as an interaction ritual chain—group interactions that involve the mutual focus of attention, as well as emotional entrainment in the form of bodily synchronization and mutual arousal. Such interactions imbue group members with emotional energy and strengthen group attachment.

We suspect that group members who identify strongly with a group will generally exhibit low levels of discursive diversity, reflecting their shared understandings of the world. However, when they are engaged in interaction ritual chains that are characterized by high levels of discursive diversity, we anticipate that group members will experience heightened emotional entrainment and thus feel even more strongly identified with the group. Moreover, although Collins initially described this process unfolding in face-to-face interactions such as spectators at sporting events or employees taking smoking breaks at work, more recent work shows that interaction ritual chains can also unfold in online settings such as employees of a multinational company discussing its core values and beliefs (DiMaggio et al. 2018). Thus, it may be possible to apply our approach to measuring discursive diversity to corpora of interactional language use among various types of groups and forecast when interactions will solidify group boundaries versus cause them to become more porous. Applications of such an approach range from predicting when political polarization will intensify to forecasting when organizational faultlines are likely to emerge.

4.3. Teams and Organizational Ambidexterity

Finally, the construct of discursive diversity can shed new light on the role that teams play in helping the organizations to which they belong navigate the tensions of exploration and exploitation (March 1991). In particular, a prominent line of research has examined the antecedents of organizational ambidexterity—the ability to simultaneously pursue both incremental and disruptive innovation (Tushman and O'Reilly 1996). Prior work has highlighted three pathways by which organizations can achieve ambidexterity: sequential, or oscillating back and forth between periods of exploration and exploitation; structural, or pursuing both objectives simultaneously by separating the two sets of activities and the capabilities needed to execute them into distinct organizational subunits; and contextual, or creating a context in which individuals can exercise appropriate judgment about when to pursue exploration versus exploitation (O'Reilly and Tushman 2013, Rogan and Mors 2014).

Research on the role of teams in fostering organizational ambidexterity has tended to focus on the structural pathway and highlighted the role of top management teams in balancing the competing interests of the units dedicated to building new capabilities versus those focused on harvesting existing ones (Jansen et al. 2008, Carmeli and Halevi 2009, Mihalache et al. 2014). With the introduction of discursive diversity, we open an alternative route to organizational ambidexterity: equipping teams throughout the organization—not just at the top—with the skills needed to match their levels of discursive diversity to their task requirements. In a sense, organizations possessing such teams can be thought of as integrating the sequential—oscillating between high and low levels of discursive diversity—and contextual—matching those oscillations to cycles of exploration versus exploitation—pathways to ambidexterity.

4.4. Limitations and Future Directions

Given the data available to us, this study has certain limitations, which point to useful avenues for future research. First, although we relied on qualitative interviews with project managers and platform leaders and qualitatively coded transcripts of Slack communications, we ultimately had to rely on time-based proxies for ideational versus coordination tasks. Whereas the teams in our data uniformly followed the Scrum framework, in other settings teams vary in how they structure their work. To generalize our approach to other types of teams, one would ideally pair electronic communications data with other artifacts that reveal the nature of tasks the team is engaged in at different points in time. In the context of software development, for example, researchers could potentially tap into teams' Github repositories to inspect how the nature of coding tasks shifts over time.

Although our outcome measure—timely delivery of milestone objectives—was economically consequential in our setting, we acknowledge that it is nevertheless a crude indicator of team success. Moreover, we only had access to communication in public Slack channels and therefore could not measure the consequences of discursive diversity expressed in private, direct messages between team members. Future research would benefit from having access to both public and private communication and from pairing objective indicators of success (e.g., timeliness, productivity, and profitability) with subjective measures of team learning, relationship quality, and well-being to build a fuller account of the outcomes that are shaped by discursive diversity.

Despite our use of models that include team fixed effects and thus account for unobserved, time-invariant heterogeneity between teams, we are not able to make strong causal claims with our empirical setup. For example, it is conceivable that teams that are confident they will achieve milestone success might feel more free to communicate in discursively diverse ways, whereas teams that know they will miss a milestone might narrow their ambitions and communicate with less discursive diversity. Given that any potential shifts in discursive diversity as response to anticipated performance are most likely to occur in later milestones and during longer milestones, the inclusion of milestone number and milestone length fixed effects in our models partially addresses this concern. To fully overcome the threat of reverse causality, we suspect that it will be necessary to design interventions that nudge teams to communicate in more or less discursively diverse ways, randomly assign teams to receive these different treatments, and then compare the performance of the two sets of teams (Hauser et al. 2018).

Finally, our findings raise a number of open questions that we leave to future research. For example, to what extent are teams aware of shifts in the task requirements they face—particularly when they are not operating by relatively fixed process such as Scrum? What practices lead teams to increase or decrease their levels of cognitive diversity? What is the role of the team leader in inducing such shifts?

4.5. Conclusion

This study paves the way for further work that uncovers the linguistic manifestations of shared cognition in groups. It suggests that the presumed performance tradeoff of group cognitive diversity can be overcome when teams and their leaders understand how to time their differences—encouraging members to express a breadth of meanings when the team engages in ideational tasks and, conversely, subtly tamping down discursive diversity when the team engages in coordination tasks. Teams that develop this capacity for timing differences can potentially achieve both high levels of creativity and seamless coordination.

References

Aggarwal, Ishani, Anita Williams Woolley. 2019. Team creativity, cognition, and cognitive style diversity. *Management Science* **65**(4) 1586–1599.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, Romain Wacziarg. 2003. Fractionalization. *Journal of Economic Growth* **8**(2) 155–194. doi:10.1023/A:1024471506938.

Amabile, Teresa M., Regina Conti, Heather Coon, Jeffrey Lazenby, Michael Herron. 1996. Assessing the Work Environment for Creativity. *Academy of Management Journal* **39**(5) 1154–1184. doi:10.5465/256995.

Ballard, Dawna I, Franziska Tschan, Mary J Waller. 2008. All in the timing: Considering time at multiple stages of group research. *Small Group Research* **39**(3) 328–351.

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

Berger, Peter L., Thomas Luckmann. 1967. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Anchor, New York.

Bernstein, Basil B. 1971. *Class, Codes and Control: Theoretical Studies towards a Sociology of Language*. Routledge & Kegan Paul.

Bernstein, Ethan, Jesse Shore, David Lazer. 2018. How intermittent breaks in interaction improve collective intelligence. *Proceedings of the National Academy of Sciences* **115**(35) 8734–8739. doi:10.1073/pnas.1802407115.

Bernstein, Shai, Rebecca Diamond, Timothy McQuade, Beatriz Pousada. 2019. The Contribution of High-Skilled Immigrants to Innovation in the United States 83.

Blei, David M., Andrew Y. Ng, Michael I. Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* **3**(Jan) 993–1022.

Bourgeois, L. J. 1985. Strategic Goals, Perceived Uncertainty, and Economic Performance in Volatile Environments. *Academy of Management Journal* **28**(3) 548–573. doi:10.5465/256113.

Carmeli, Abraham, Meyrav Yitzack Halevi. 2009. How top management team behavioral integration and behavioral complexity enable organizational ambidexterity: The moderating role of contextual ambidexterity. *The Leadership Quarterly* **20**(2) 207–218.

Carnabuci, Gianluca, Bálint Diószegi. 2015. Social networks, cognitive style, and innovative performance: A contingency perspective. *Academy of Management Journal* **58**(3) 881–905.

Christianson, Marlys K. 2019. More and less effective updating: The role of trajectory management in making sense again. *Administrative Science Quarterly* **64**(1) 45–86. doi:10.1177/0001839217750856. URL `https://doi.org/10.1177/0001839217750856`.

Cohen, Wesley M., Daniel A. Levinthal. 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly* **35**(1) 128–152. doi:10.2307/2393553.

Collins, Randall. 2014. *Interaction Ritual Chains*. Princeton University Press.

Converse, Sharolyn, JA Cannon-Bowers, E Salas. 1993. Shared mental models in expert team decision making. *Individual and Group Decision Making: Current Issues*, vol. 221. Hillsdale, 221–46.

Cooke, Nancy J., Jamie C. Gorman, Christopher W. Myers, Jasmine L. Duran. 2013. Interactive Team Cognition. *Cognitive Science* **37**(2) 255–285. doi:10.1111/cogs.12009.

Corritore, Matthew, Amir Goldberg, Sameer B. Srivastava. 2019. Duality in Diversity: How Intrapersonal and Interpersonal Cultural Heterogeneity Relate to Firm Performance. *Administrative Science Quarterly* 0001839219844175doi:10.1177/0001839219844175.

Cronin, Matthew A, Laurie R Weingart, Gergana Todorova. 2011. Dynamics in groups: Are we there yet? *The academy of management annals* **5**(1) 571–612.

Cropley, Arthur. 2006. In praise of convergent thinking. *Creativity research journal* **18**(3) 391–404.

Danescu-Niculescu-Mizil, Cristian, Lillian Lee, Bo Pang, Jon Kleinberg. 2012. Echoes of power: Language effects and power differences in social interaction. *Proceedings of the 21st International Conference on World Wide Web*. WWW '12, Association for Computing Machinery, Lyon, France, 699–708. doi:10.1145/2187836.2187931.

Davison, Graydon, Deborah Blackman. 2005. The role of mental models in innovative teams. *European Journal of Innovation Management* **8**(4) 409–423. doi:10.1108/14601060510627795.

de Vaan, Mathijs, Balazs Vedres, David Stark. 2015. Game Changer: The Topology of Creativity. *American Journal of Sociology* **120**(4) 1144–1194. doi:10.1086/681213.

de Woot, Philippe, Hubert Heyvaert, François Martou, Camilla de Crombeke. 1977. Strategic Management: An Empirical Study of 168 Belgian Firms. *International Studies of Management & Organization* **7**(3/4) 60–75.

Dess, Gregory G. 1987. Consensus on strategy formulation and organizational performance: Competitors in a fragmented industry. *Strategic Management Journal* **8**(3) 259–277. doi:10.1002/smj.4250080305.

Detert, James R., Amy C. Edmondson. 2011. Implicit Voice Theories: Taken-for-Granted Rules of Self-Censorship at Work. *Academy of Management Journal* **54**(3) 461–488. doi:10.5465/amj.2011.61967925.

DiMaggio, Paul, Clark Bernier, Charles Heckscher, David Mimno. 2018. Interaction ritual threads: Does IRC theory apply online? *Ritual, Emotion, Violence*. Routledge, 99–142.

Doyle, Gabriel, Amir Goldberg, Sameer Srivastava, Michael C Frank. 2017. Alignment at work: Using language to distinguish the internalization and self-regulation components of cultural fit in organizations. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 603–612.

Eliasoph, Nina, Paul Lichterman. 2003. Culture in Interaction. *American Journal of Sociology* **108**(4) 735–794. doi:10.1086/367920.

Fiol, C. Marlene. 1994. Consensus, Diversity, and Learning in Organizations. *Organization Science* **5**(3) 403–420. doi:10.1287/orsc.5.3.403.

Fiol, C. Marlene, Marjorie A. Lyles. 1985. Organizational Learning. *Academy of Management Review* **10**(4) 803–813. doi:10.5465/amr.1985.4279103.

Fiore, Stephen M., Eduardo Salas. 2004. Why we need team cognition. *Team Cognition: Understanding the Factors That Drive Process and Performance*. American Psychological Association, Washington, DC, US, 235–248. doi:10.1037/10690-011.

Fleishman, Edwin A., Stephen J. Zaccaro. 1992. Toward a taxonomy of team performance functions. *Teams: Their Training and Performance*. Ablex Publishing, Westport, CT, US, 31–56.

Foucault, Michel. 2002. *Archaeology of Knowledge*. Psychology Press.

Garg, Nikhil, Londa Schiebinger, Dan Jurafsky, James Zou. 2018. Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences* **115**(16) E3635–E3644. doi:10.1073/pnas.1720347115.

Gersick, Connie JG. 1991. Revolutionary change theories: A multilevel exploration of the punctuated equilibrium paradigm. *Academy of management review* **16**(1) 10–36.

Gibbs, Jack P, Walter T Martin. 1962. Urbanization, technology, and the division of labor: International patterns. *American sociological review* 667–677.

Gibson, Cristina, Freek Vermeulen. 2003. A healthy divide: Subgroups as a stimulus for team learning behavior. *Administrative science quarterly* **48**(2) 202–239.

Goffman, Erving. 1959. *The Presentation of Self in Everyday Life*. Anchor.

Goldberg, Amir, Sameer B. Srivastava, V. Govind Manian, William Monroe, Christopher Potts. 2016. Fitting In or Standing Out? The Tradeoffs of Structural and Cultural Embeddedness. *American Sociological Review* **81**(6) 1190–1222. doi:10.1177/0003122416671873.

Gumperz, John J. 1982. Discourse Strategies. /core/books/discourse-strategies/46D4D801BCC1FF7E6E31B2A19E45E92B. doi:10.1017/CBO9780511611834.

Hannan, Michael T., Gaël Le Mens, Greta Hsu, Balázs Kovács, Giacomo Negro, László Pólos, Elizabeth Pontikes, Amanda J. Sharkey. 2019. *Concepts and Categories: Foundations for Sociological and Cultural Analysis*. Columbia University Press.

Hargadon, Andrew B., Beth A. Bechky. 2006. When Collections of Creatives Become Creative Collectives: A Field Study of Problem Solving at Work. *Organization Science* **17**(4) 484–500. doi:10.1287/orsc.1060.0200.

Hauser, Oliver P, Francesca Gino, Michael I Norton. 2018. Budging beliefs, nudging behaviour. *Mind & Society* **17**(1-2) 15–26.

Healey, Mark P., Timo Vuori, Gerard P. Hodgkinson. 2015. When Teams Agree While Disagreeing: Reflexion and Reflection in Shared Cognition. *Academy of Management Review* **40**(3) 399–422. doi:10.5465/amr.2013.0154.

Hinds, Pamela J., Diane E. Bailey. 2003. Out of Sight, Out of Sync: Understanding Conflict in Distributed Teams. *Organization Science* **14**(6) 615–632. doi:10.1287/orsc.14.6.615.24872.

Hong, Lu, Scott E. Page. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences* **101**(46) 16385–16389. doi:10.1073/pnas.0403723101.

Hrebiniak, Lawrence G., Charles C. Snow. 1982. Top-Management Agreement and Organizational Performance. *Human Relations* **35**(12) 1139–1157. doi:10.1177/001872678203501205.

Huber, George P. 1991. Organizational Learning: The Contributing Processes and the Literatures. *Organization Science* **2**(1) 88–115. doi:10.1287/orsc.2.1.88.

Ireland, Molly E., Richard B. Slatcher, Paul W. Eastwick, Lauren E. Scissors, Eli J. Finkel, James W. Pennebaker. 2011. Language Style Matching Predicts Relationship Initiation and Stability. *Psychological Science* **22**(1) 39–44. doi:10.1177/0956797610392928.

Jackson, Susan E., Jane E. Dutton. 1988. Discerning Threats and Opportunities. *Administrative Science Quarterly* **33**(3) 370–387. doi:10.2307/2392714.

Janis, Irving L. 1971. Groupthink. *Psychology today* **5**(6) 43–46.

Jansen, Justin JP, Gerard George, Frans AJ Van den Bosch, Henk W Volberda. 2008. Senior team attributes and organizational ambidexterity: The moderating role of transformational leadership. *Journal of Management Studies* **45**(5) 982–1007.

Jehn, Karen A, Gregory B Northcraft, Margaret A Neale. 1999. Why differences make a difference: A field study of diversity, conflict and performance in workgroups. *Administrative science quarterly* **44**(4) 741–763.

Kahneman, Daniel. 1991. Article Commentary: Judgment and Decision Making: A Personal View. *Psychological Science* **2**(3) 142–145. doi:10.1111/j.1467-9280.1991.tb00121.x.

Kelly, Janice R, Joseph E McGrath. 1985. Effects of time limits and task types on task performance and interaction of four-person groups. *Journal of personality and social psychology* **49**(2) 395.

Kilduff, Martin, Reinhard Angelmar, Ajay Mehra. 2000. Top Management-Team Diversity and Firm Performance: Examining the Role of Cognitions. *Organization Science* **11**(1) 21–34. doi:10.1287/orsc.11.1.21.12569.

Klimoski, Richard, Susan Mohammed. 1994. Team mental model: Construct or metaphor? *Journal of Management* **20**(2) 403–437. doi:10.1016/0149-2063(94)90021-3.

Knight, Don, Craig L Pearce, Ken G Smith, Judy D Olian, Henry P Sims, Ken A Smith, Patrick Flood. 1999. Top management team diversity, group process, and strategic consensus. *Strategic Management Journal* **20**(5) 445–465.

Knorr Cetina, Karin, Urs Bruegger. 2002. Global Microstructures: The Virtual Societies of Financial Markets. *American Journal of Sociology* **107**(4) 905–950. doi:10.1086/341045.

Kozlowski, Austin C., Matt Taddy, James A. Evans. 2019. The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings. *American Sociological Review* **84**(5) 905–949. doi:10.1177/0003122419877135.

Lewis, Lewis. 2002. *Convention*. Blackwell, Oxford.

March, James G. 1991. Exploration and Exploitation in Organizational Learning. *Organization Science* **2**(1) 71–87. doi:10.1287/orsc.2.1.71.

Marks, Michelle A., John E. Mathieu, Stephen J. Zaccaro. 2001. A Temporally Based Framework and Taxonomy of Team Processes. *Academy of Management Review* **26**(3) 356–376. doi:10.5465/amr.2001.4845785.

Maznevski, Martha L., Katherine M. Chudoba. 2000. Bridging Space Over Time: Global Virtual Team Dynamics and Effectiveness. *Organization Science* **11**(5) 473–492. doi:10.1287/orsc.11.5.473.15200.

McGrath, Joseph E. 1991. Time, Interaction, and Performance (TIP): A Theory of Groups. *Small Group Research* **22**(2) 147–174. doi:10.1177/1046496491222001.

Mihalache, Oli R, Justin JP Jansen, Frans AJ Van den Bosch, Henk W Volberda. 2014. Top management team shared leadership and organizational ambidexterity: A moderated mediation framework. *Strategic Entrepreneurship Journal* **8**(2) 128–148.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, Jeff Dean. 2013. Distributed Representations of Words and Phrases and their Compositionality. C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, K. Q. Weinberger, eds., *Advances in Neural Information Processing Systems 26*. Curran Associates, Inc., 3111–3119.

Milliken, Frances J, Luis L Martins. 1996. Searching for common threads: Understanding the multiple effects of diversity in organizational groups. *Academy of management review* **21**(2) 402–433.

Mobasseri, Sanaz, Amir Goldberg, Sameer B Srivastava. 2019. What is cultural fit? *The Oxford Handbook of Cognitive Sociology* 305.

Mohammed, Susan, Lori Ferzandi, Katherine Hamilton. 2010. Metaphor No More: A 15-Year Review of the Team Mental Model Construct. *Journal of Management* **36**(4) 876–910. doi:10.1177/0149206309356804.

Mohammed, Susan, Richard Klimoski, Joan R. Rentsch. 2000. The Measurement of Team Mental Models: We Have No Shared Schema. *Organizational Research Methods* **3**(2) 123–165. doi:10.1177/109442810032001.

Moore, Don A., Lloyd Tanlu, Max H. Bazerman. 2010. Conflict of interest and the intrusion of bias. *Judgment and Decision Making* **5**(1) 37–53.

Niederhoffer, Kate G., James W. Pennebaker. 2002. Linguistic Style Matching in Social Interaction. *Journal of Language and Social Psychology* **21**(4) 337–360. doi:10.1177/026192702237953.

O'Reilly, Charles A, Michael L Tushman. 2013. Organizational ambidexterity: Past, present, and future. *Academy of management Perspectives* **27**(4) 324–338.

Page, Scott E. 2008. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies - New Edition*. Princeton University Press.

Pelled, Lisa Hope, Kathleen M. Eisenhardt, Katherine R. Xin. 1999. Exploring the Black Box: An Analysis of Work Group Diversity, Conflict and Performance. *Administrative Science Quarterly* **44**(1) 1–28. doi:10.2307/2667029.

Porter, Shanette C., Michelle Rheinschmidt-Same, Jennifer A. Richeson. 2016. Inferring Identity From Language: Linguistic Intergroup Bias Informs Social Categorization. *Psychological Science* **27**(1) 94–102. doi:10.1177/0956797615612202.

Prince, Carolyn, Eduardo Salas. 1993. Training and research for teamwork in the military aircrew. *Cockpit Resource Management*. Academic Press, San Diego, CA, US, 337–366.

Reagans, Ray, Ella Miron-Spektor, Linda Argote. 2016. Knowledge utilization, coordination, and team performance. *Organization Science* **27**(5) 1108–1124.

Ren, Yuqing, Linda Argote. 2011. Transactive memory systems 1985–2010: An integrative framework of key dimensions, antecedents, and consequences. *Academy of Management Annals* **5**(1) 189–229.

Rickford, John R., Greg J. Duncan, Lisa A. Gennetian, Ray Yun Gou, Rebecca Greene, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, Lisa Sanbonmatsu, Andres E. Sanchez-Ordoñez, Matthew Sciandra, Ewart Thomas, Jens Ludwig. 2015. Neighborhood effects on use of African-American Vernacular English. *Proceedings of the National Academy of Sciences* **112**(38) 11817–11822. doi:10.1073/pnas.1500176112.

Rogan, Michelle, Marie Louise Mors. 2014. A network perspective on individual-level ambidexterity in organizations. *Organization Science* **25**(6) 1860–1877.

Samila, Sampsa, Olav Sorenson. 2017. Community and Capital in Entrepreneurship and Economic Growth. *American Sociological Review* **82**(4) 770–795. doi:10.1177/0003122417711699.

Schwaber, Ken. 1997. SCRUM Development Process. Jeff Sutherland, Cory Casanave, Joaquin Miller, Philip Patel, Glenn Hollowell, eds., *Business Object Design and Implementation*. Springer, London, 117–134. doi:10.1007/978-1-4471-0947-1_11.

Sørensen, Jesper B. 2002. The Strength of Corporate Culture and the Reliability of Firm Performance. *Administrative Science Quarterly* **47**(1) 70–91. doi:10.2307/3094891.

Srikanth, Kannan, Sarah Harvey, Randall Peterson. 2016. A dynamic perspective on diverse teams: Moving from the dual-process model to a dynamic coordination-based model of diverse team performance. *Academy of Management Annals* **10**(1) 453–493.

Srivastava, Sameer B, Mahzarin R Banaji. 2011. Culture, cognition, and collaborative networks in organizations. *American Sociological Review* **76**(2) 207–233.

Srivastava, Sameer B., Amir Goldberg, V. Govind Manian, Christopher Potts. 2018. Enculturation Trajectories: Language, Cultural Adaptation, and Individual Outcomes in Organizations. *Management Science* **64**(3) 1348–1364. doi:10.1287/mnsc.2016.2671.

Strauss, Claudia, Naomi Quinn. 1997. *A Cognitive Theory of Cultural Meaning*. Cambridge University Press.

Thompson, Leigh, Gary Alan Fine. 1999. Socially Shared Cognition, Affect, and Behavior: A Review and Integration. *Personality and Social Psychology Review* **3**(4) 278–302. doi:10.1207/s15327957pspr0304_1.

Tushman, Michael L, Charles A O'Reilly. 1996. Ambidextrous organizations: Managing evolutionary and revolutionary change. *California management review* **38**(4) 8–29.

Volk, Stefan, Matthew J Pearsall, Michael S Christian, William J Becker. 2017. Chronotype diversity in teams: Toward a theory of team energetic asynchrony. *Academy of Management Review* **42**(4) 683–702.

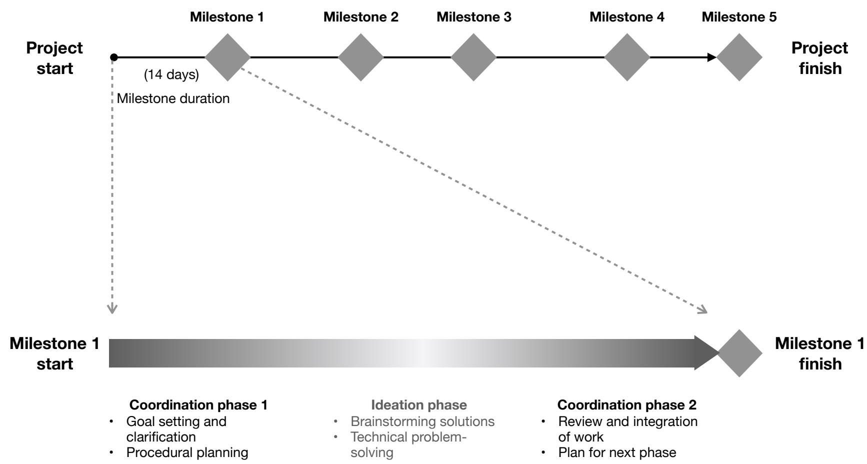
Walsh, George B. 1988. Sublime Method: Longinus on Language and Imitation. *Classical Antiquity* **7**(2) 252–269. doi:10.2307/25010890.

Wegner, Daniel M. 1987. Transactive memory: A contemporary analysis of the group mind. *Theories of group behavior*. Springer, 185–208.

Wegner, Daniel M, Ralph Erber, Paula Raymond. 1991. Transactive memory in close relationships. *Journal of personality and social psychology* **61**(6) 923.

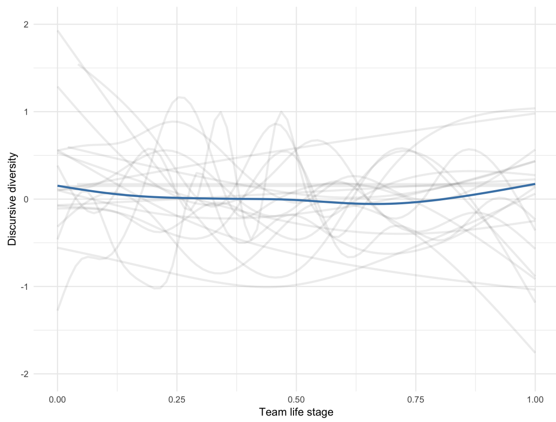
Williams, K, Charles O'Reilly. 1998. The complexity of diversity: A review of forty years of research. *Research in organizational behavior* **21** 77–140.

FIGURES

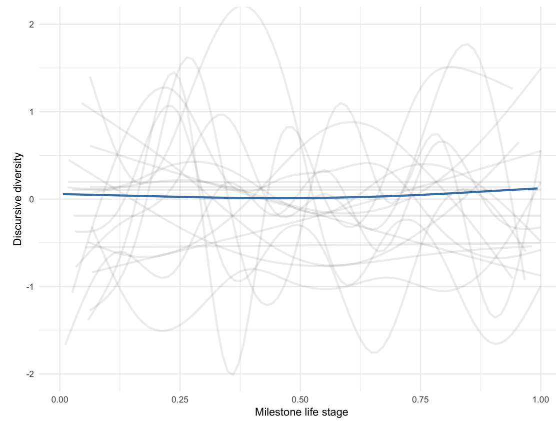


Note: The lower part of the figure depicts the three distinct phases that occur within each milestone, in this case Milestone 1.

Figure 1 Schematic Representation of a Hypothetical Project that Includes Five Milestones of Varying Length



(a) Discursive diversity by team life stage for 20 randomly selected teams



(b) Discursive diversity by milestone life stage for 20 randomly selected milestone phases

Note: The blue lines, respectively, represent the mean level of discursive diversity across all 117 teams at a given team life stage (Panel a) and the mean level of discursive diversity across all milestone phases for a given milestone life stage (Panel b).

Figure 2 Discursive Diversity Across Team and Milestone Life Stages

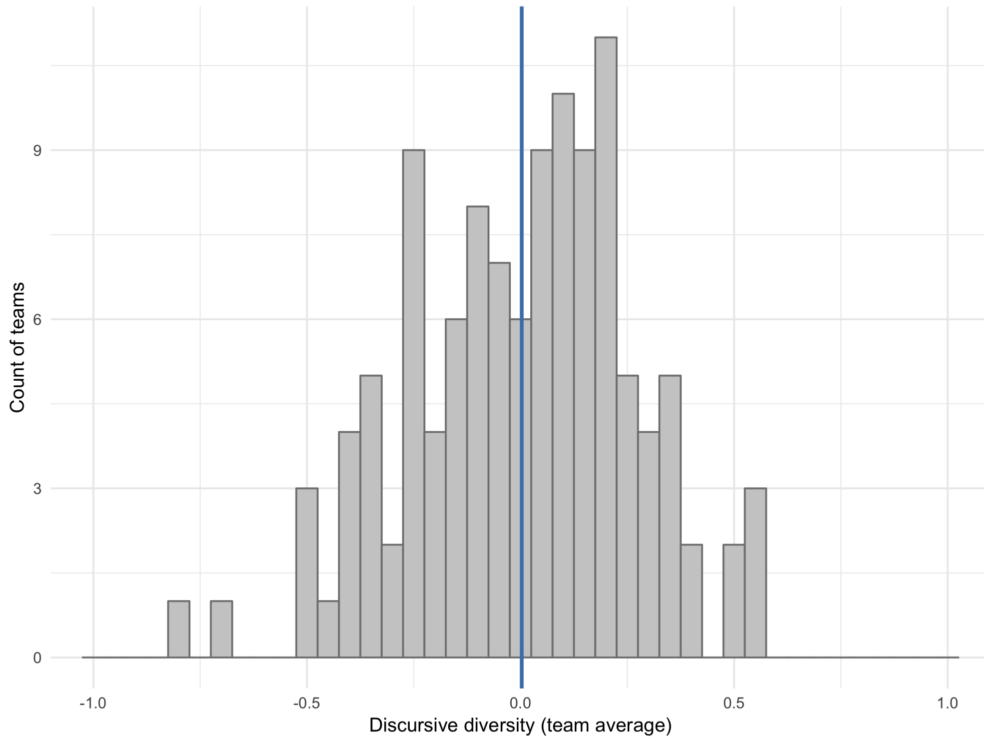


Figure 3 **Distribution of Discursive Diversity across 117 teams**

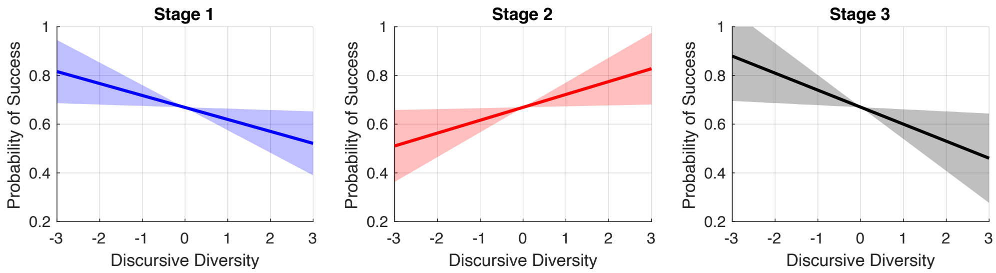


Figure 4 **Marginal effects of Discursive Diversity on Milestone Success (Model 4, Table 3)**

TABLES

Table 1 Topics identified from teams' conversations and illustrative key words

Topic label	Illustrative key words
Project management	milestone, project, week, scope, channel, team, review, help, feature
Programming issues	data, code, file, server, error, page, use, mean, search
Design	design, screen, page, image, button, icon, view, great
User interface considerations	build, screen, push, notification, message, account, challenge, profile
Software testing	(Trello) card, test, admin, testing, error, data, staging, worker, message
Payments and transactions	account, card, stripe, payment, page, info, customer, charge
Backend infrastructure issues	endpoint, token, login, request, error, return, server, response, session, backend
Hardware and devices	screen, button, setting, build, phone, keyboard, dashboard, store, text
Delivery functionality in apps	post, status, restaurant, meal, content, code, submission, image
Multimedia functionality in apps	video, audio, player, play, artist, song, playlist, track, music
Contracts and negotiations	contract, deal, email, calculation, document, template, project, talent
Purchasing functionality in apps	checkout, price, order, vendor, item, location, delivery, referral

Note: Topics are shown in order of relative prevalence across teams' conversations.

Table 2 Analogy Tasks Based on Word Embedding Model Trained on Entire Set of Teams' Slack Archives

Analogy task	**Answer**
Bug - code = ?	Issue
Milestone + deliverable = ?	Sprint
Sprint - pressure = ?	Phase
Man + casual = ?	Dude
Instagram - photos = ?	Facebook
Machine - software = ?	Device
Machine + intelligent = ?	Brain
California - startup = ?	Australia
Human + desires + art = ?	Culture
Visual - creative = ?	Polish
Team - community = ?	@-tag
Man + programmer = ?	Beard
Woman + programmer = ?	Roadblock

Table 3: **Illustrative Quotes from Team Conversations From Milestone Stages with Varying Levels of Discursive Diversity**

Milestone stage	Illustrative quotes from team's Slack conversation	Discursive diversity (standardized)
Stage 1	[Team receives a new feature request from CLIENT:] Engineer 1: @[PM]: Here is the document outlining the data problem with all [APP OUTPUT] being sent to device [LINK]. It's not that trivial, this preset list - it's quite a lot of work and each mapping corresponds to a different set of [FEATURE] rules. I just don't think I'll have it done by tomorrow. [. . .] [CLIENT] sounds pretty flexible in that correspondence. Once they decide about these time restrictions I think we could have this whole thing wrapped up and finished by next week? PM: Okay. [CLIENT] has just sprung it on us, so I'll just let [CLIENT] know. Engineer 1: Yeah, I'm sure it won't be problem as it's another last-minute feature that wasn't planned. [Next day:] PM: Morning Team, how are we doing? Engineer 2: Any feedback about the document? PM: No, nothing yet. Engineer 2: Maybe by today. PM: Fingers crossed. Engineer 1: Do we have a delivery date for the next milestone? I'd like to get everything done and wrapped sooner rather than later. I'm finishing up [TO DO] and then I need to remove the ability to [APP FUNCTIONALITY] Engineer 2: Will that still take care of [APP FUNCTIONALITY]? Engineer 1: @Engineer 2: Yes, it will include everything we currently show, plus any [FEATURES] in the future PM: Yes, our delivery date is [DATE].	-1.00283612

Stage 2	Engineer 1: @Channel: Hi ladies - I 've just pushed a big update. We now should have [FEATURES] shown as per the client's request. I have implemented the preset [FEATURES], as well as the ability to repeat by hour or day. I've also changed the delete buttons so that they are timed - you have to hold them down for a second for the delete action to execute - I thought this was better than a confirmation dialogue, especially when using touch screens, and it looks pretty slick. I'm going to do some more extensive testing tomorrow, but I think that's all of [CLIENT]'s feedback done on my side. PM: Awesome @Engineer 1. Have you pushed the changes? Engineer 1: Yes, they are up on [CODE PLATFORM]. Let 's test it ourselves properly before we give it to [CLIENT] to test. PM: Yup. Okay give me an hour. I 'll go through it. PM: Hey @Engineer 1, I'm holding your build to ransom again. So don't share any builds with client until I say so [EMOJI] Engineer 2: No sending, no way! PM: @Engineer 2: Remember we said we were going to have a page for historic [FEATURES] and the ability to export to csv? Engineer 1: You're kidding. I mean we mentioned reporting but that was never included in a list of the feedback or feature requests. PM: lol no that was a discussion between me and you. Engineer 1: What exactly do you want to be able to export? PM: I'm actually thinking of exporting [LIST OF FEATURES] [...] Engineer 1: In any case, that's not really the goal of the app, or is it? It's to make sure [FUNCTIONALITY], not to give analytics on [ITEM] performance. I could make a quick page that just lists all incomplete [FEATURE]s?	0.89299718

Stage 3	[Engineer 1 discovers a bug in the team's code and has raised the issue to the team.] PM: To be honest, I think this is a problem if the person setting [FEATURE] and the person receiving [FEATURE] are in different time zones. Engineer 1: I know what the problem is, will fix it asap. PM: Try setting a [FEATURE] with a device that is in a different time zone than the server. Engineer 1: Yeah, that's what I thought. I think the solution is to remove time zone info from the data we send to the server. So, time is always just a string and it will show the same regardless of where you are. PM: Okay, that works. Engineer 1: Cause there might have been some automatic conversion happening. PM: Yeah, I agree. Engineer 1: Great. Will let you know once I've fixed these things tonight.	-1.25892899

Table 4 Summary Statistics and Bivariate Correlations

	Summary statistics (N = 450)		Bivariate correlations						
	Mean	**St. Dev.**	**Success**	**Discursive diversity (st.)**	**Topical diversity (st.)**	**Team size**	**Role diversity**	**Gender diversity**	**Country diversity**
Success	0.662	0.473							
Discursive diversity (st.)	-0.035	0.829	-0.03 .						
Topical diversity (st.)	0.001	1	0.03 .	-0.04 .					
Team size	9.076	4.578	0.04 .	0 .	0.13 **				
Role diversity	0.763	0.099	0.15 **	0.11 *	0.1 *	0.58 ***			
Gender diversity	0.404	0.244	-0.05 .	0.01 .	0.01 .	0.04 .	-0.13 **		
Country diversity	0.62	0.179	0.04 .	0.13 **	0.05 .	0.18 ***	0.25 ***	0.14 **	
Milestone duration (days)	14.024	18.242	0.02 .	-0.04 .	0.02 .	0.05 .	-0.01 .	0.04 .	-0.07 .

Table 5 Linear Probability Models of Milestone Success on Covariates

	Dependent variable: Milestone success			
	Model			
	(1)	(2)	(3)	(4)
Topic Diversity	0.00586	0.00525	-0.000388	-0.00234
	(0.24)	(0.21)	(-0.01)	(-0.09)
Discursive Diversity (mean)		-0.0299		
		(-0.89)		
Discursive Diversity (stage 1)			-0.0540*	-0.0491*
			(-2.42)	(-2.21)
Discursive Diversity (stage 2)			0.0499*	0.0528*
			(2.00)	(2.11)
Discursive Diversity (stage 3)			-0.0622*	-0.0699*
			(-1.99)	(-2.24)
Constant	0.668***	0.668***	0.668***	0.669***
	(2108.90)	(1661.30)	(1526.75)	(1481.75)
N	509	509	487	487
Team Fixed Effects	Yes	Yes	Yes	Yes
Milestone Length Fixed Effects	Yes	Yes	Yes	Yes
Milestone Number Fixed Effects	No	No	No	Yes

Table 6 Conditional Logit Models of Milestone Success on Covariates

	Dependent variable: Milestone success			
	Model			
	(1)	(2)	(3)	(4)
Topic Diversity	-0.007 (0.045)	-0.011 (0.045)	-0.023 (0.049)	-0.027 (0.05)
Discursive Diversity (mean)		-0.100 (0.006)		
Discursive Diversity (stage 1)			-0.131** (0.044)	-0.113* (0.048)
Discursive Diversity (stage 2)			0.109* (0.048)	0.119* (0.05)
Discursive Diversity (stage 3)			-0.151** (0.054)	-0.168** (0.055)
N	509	509	487	487
Team Fixed Effects	Yes	Yes	Yes	Yes
Milestone Length Fixed Effects	Yes	Yes	Yes	Yes
Milestone Number Fixed Effects	No	No	No	Yes