The Grammar of Diverse Knowledge Combination in Teams:
Diverse Recombination vs. Division of Specialized Labor vs. Social Processes

Draft: 01/10/24

Kevin J. Boudreau (Northeastern University & NBER)
Nilam Kaushik (IIM Bangalore)

Abstract:

How does a team's knowledge diversity affect innovation performance? Diversity of knowledge (training, expertise) is often suggested to improve innovation performance. Yet, theory remains fragmented, and discerning causal evidence is limited. We delineate three distinct theoretical perspectives on this question that have largely been emphasized in separate literatures: Diverse Recombination, the Division of Specialized Labor, and Social Processes. We study the existence, workings, and interactions of associated mechanisms in a field experiment. 872 adults from Business, Humanities, Computer Science, Design, Engineering, Health Sciences, Humanities, Law, and Sciences were randomly assigned to 218 teams of 4 to work in a 3-week innovation sprint. We first show intermediate diversity led to the highest quality and relative novelty of innovations, while novelty conditional on quality monotonically increased with diversity. Our main analysis then studies underlying mechanisms. We document experimental evidence of the (co)existence of the 3 sets of mechanisms—and that promoting any one set of mechanisms tends to be antagonistic to the others, implying sharp tradeoffs in designing team knowledge. In this particular context, Social Processes dominated, where knowledge subgroups and intra-group faultlines were most important in shaping innovation. Most effective were "balanced teams" of 2-and-2 from distinct fields, so long as they came from fields related to the problem at hand. (We also document experimental evidence of similar effects in gender-balanced teams.) The results show that the social embeddedness of knowledge can be as important as the role of knowledge-as-input to the innovation process.

Keywords: Knowledge, Diversity, Innovation, Teams, Product Development, Platforms, Field Experiment

Boudreau: (corresponding author) Northeastern University and National Bureau of Economic Research (e-mail: kevin@northeastern.edu). Kaushik: Indian Institute of Management Bangalore (e-mail: nilam.kaushik@iimb.ac.in). We thank members of supporting organizations and particularly Maria Costa De Sousa, Mavez Dabas, Sylvain Demortier, Anthony Donaldson, Austen Keene, Afan Khan, Atif Khan, Abhinav Kharbanda, Dyan Khor, Raghavi Kirouchenaradjou, Sreerag Sreenath Mandakathil, Patrick McGrath, Robert Hughes, Michael Orr, Olga Ozhereleva, Kaushik Padmanabhan, Edwige Poinssot, Prathamesh Kisan Tajane, Nikin Tharan, Maureen Underhill, Robert Whelan, and Katherine Wilhoit. For their especially helpful comments on this study, we thank Sekou Bermiss, Dan Elfenbein, Roberto Fernandez, Lee Fleming, Anne Marie Knott, Ben Jones, Liz Lyons, Matt Marx, Siobhan O'Mahony, Oren Reshef, Florent Teodoridis, Keyvan Vakili, Freek Vermeulen, and seminar participants at Boston University, UNC Chapel Hill, Washington University, and the Academy of Management. We would also like to acknowledge generous financial support from the Kauffman Foundation and Northeastern DMSB. The views expressed here are those of the authors alone and do not necessarily reflect those of supporting institutions, funders or supporters involved with this study. All errors are our own.

1 Introduction

The study of how diversity of team member characteristics affects team performance has grown considerably since the late 1980s, with the bulk of research focusing on how demographic and task-related differences relate to team processes across a wide range of contexts (Roberson 2019).[1] The current study focuses on a specific area where there has yet been less accumulated research, where theoretical views remain fragmented, and where there are still few clear-cut answers: How does diversity (heterogeneity, differences) of knowledge (training, expertise) affect the performance of innovation teams?[2] In this study, we delineate and investigate the role of three distinct theoretical perspectives that have been emphasized in separate literature explaining and predicting the link between knowledge diversity and innovation performance (Section 2). Two of these perspectives emphasize knowledge differences as an input to innovation processes, whereas the third perspective emphasizes knowledge differences as shaping team innovation processes.

The "Diverse Recombination" perspective (Section 2.1) predicts that diverse knowledge held by problem-solvers enables the synthesis of novel and potentially high-quality breakthrough innovations—particularly from atypical knowledge combinations (e.g., Fleming 2001; Uzzi et al. 2013). Here, diversity is understood to expand the innovative search for new ideas. By contrast, the "Division of Specialized Labor" perspective (Section 2.2) emphasizes the substantial burden of knowledge that must be overcome to develop new innovative solutions, with the view that doing so is best accomplished by assembling specialists whose knowledge maps to the nature and decomposability of the innovation problem (e.g., Jones 2009; Simon 1991). Notwithstanding compelling seminal theory and the demonstration of supporting empirical relationships supporting these perspectives—particularly on patented technologies and published academic papers (Section 2)—these two views of knowledge-as-an-input considerably diverge on their implications. Moreover, important questions linger on precisely how knowledge inputs translate to innovation outputs, including questions regarding benefits of breadth-versus-depth of knowledge in generating breakthroughs (e.g., Kaplan and Vakili 2015; Teodoridis, Bikard, and Vakili 2019).

The third theoretical perspective, which we refer to here as "Social Processes" (Section 2.3), encompasses a large body of work that proceeds with a highly generalized notion of diversity, based in differences of social categories and identity (Roberson 2019). Social categorical differences are theorized to shape social interactions and, thus, team processes and productivity (e.g., Pelled, Eisenhardt, and Xin 1999;

[1] Prior research has studied diversity in group decision-making and team work of various kinds in an array of characteristics, such as age (e.g., Zhang and Guler 2020), gender (e.g., Joshi 2014; Yang et al. 2022), race (Smith-Doerr, Alegria, and Sacco 2017; Bermiss, Green, and Hand 2023), task-related differences, tenure, and demography (e.g., Huckman and Staats 2011; Perretti and Negro 2007; Ancona and Caldwell 1992; Zenger and Lawrence 1989), interests (e.g., Ren, Chen, and Riedl 2016), motivations (e.g., Pollok et al. 2021), culture and beliefs (e.g., Corritore, Goldberg, and Srivastava 2020), social network position (Reagans, Zuckerman, and McEvily 2004; Brian Uzzi and Spiro 2005), and institutional affiliation. Contexts studied include senior executive teams, door to door canvassing, sports teams, sales teams, multinational consulting teams, school homework groups, political campaigns, undergrad laboratory subjects, start-up founding teams, project evaluation teams, military, and more.

[2] Demographic traits can themselves be associated with differences in knowledge at least in the sense of information possessed, use of heuristics or tools, cognitive representations, and mental models (Page 2019), or perhaps even simply be correlated with training and expertise. Our study, however, focuses on training and expertise and we will control for these other factors.

Gibson and Gibbs 2006), as when intra-team faultlines and subgroups either impede or stimulate team information-processing (Lau and Murnighan 2005; Gibson and Vermeulen 2003). These theories have largely been developed and tested in relation to demographic traits in prior studies. However, there are important reasons why we might expect differences in knowledge, training, and expertise to be similarly salient as sociological categories (Section 2.3.i).

Although there remain theoretical questions in each perspective, here we argue for integrating these largely separate theoretical perspectives—so as to better understand, predict, and guide understanding of the knowledge diversity-innovation link. We argue that the sets of mechanisms emphasized in these distinct theoretical perspectives should *jointly* co-determine the knowledge diversity-innovation link (Section 2). By taking steps towards integrating these perspectives, the current study seeks to provide greater explanatory and predictive power of the causal link between knowledge diversity on innovation performance.

There is a pressing need to address this gap in understanding how to design optimal teams. In a period in which productivity growth and R&D returns have slowed (Bloom et al. 2020), teams are emerging as the primary unit of production for technology innovation, product development, and scientific discovery (Jones 2009; Wuchty, Jones, and Uzzi 2007). Teams are also becoming a building block of modern organizations (Mathieu et al. 2013). Nonetheless, it remains difficult on the basis of existing research to offer prescriptions to innovation managers on how to best to design knowledge diversity into teams (Vakili and Kaplan 2021), short of relying on models that assume the underlying mechanisms (see discussions in Page (2019) and Hong and Page (2004)).

A key problem is the yet limited discerning literature on how knowledge leads to innovation performance, while sorting out underlying mechanisms (Section 2). The large empirical literature on team diversity of many kinds has not yet squarely addressed these questions (see Horwitz and Horwitz 2007; Jackson, Joshi, and Erhardt 2003; Joshi and Roh 2009; Roberson 2019). Nor has an exciting new stream of experimental and causal studies (e.g., Hoogendoorn, Parker, and Van Praag 2017; Lyons 2017; Mayo, Woolley, and Chow 2020) yet squarely addressed these questions (Appendix A).

Also leading to a pressing need to address these questions are the steep costs of implementing diversity. For example, a team assembled to provide divergent views may sacrifice relevant depth (Kaplan and Vakili 2015). And, of course, simply adding greater numbers of problem-solvers to achieve diversity adds to the wage bill, while aggravating group coordination. Assembling diverse teams in many cases also leads to constant reshuffling of ad hoc teams and foregoing the accretion of stable and routinized organization (Edmondson, Bohmer, and Pisano 2001; Huckman and Staats 2011; Taylor and Greve 2006). This tradeoff might be especially pronounced when harnessing high diversity from open and external innovators (Lifshitz-Assaf 2018; Davis 2016).

To make progress in discerning the existence, relative importance, and workings of the mechanisms shaping the knowledge diversity-innovation link, we designed and implemented a large field experiment on a work-from-anywhere (WFA) innovation collaboration platform. The study engaged 872 working age adults with training and expertise across a wide range of domains, including Business, Computer Science, Design, Engineering, Sciences, Law, Humanities, and Health and Nursing. Participants were assigned to 218 teams of 4 and given the same product development prompt related to developing a novel information technology product. Over a three-week innovation sprint, teams were required to submit a design that included a use case, technical architecture, and business plan components. A large panel of C-level executives scored projects according to overall innovation quality and commercial potential, along with the novelty of concepts.

Our analysis begins by first estimating the overall causal relationship between levels of diversity and innovation performance, finding that intermediate diversity leads to greater overall innovation quality and novelty than do highly homogenous or heterogeneous teams in this context. We find, too, that novelty increases with knowledge diversity when controlling for quality. Our main analysis then proceeds to assess underlying mechanisms leading to these patterns so as to derive more generalizable insights. We do so by contrasting the predictions of each of the theoretical perspectives in how knowledge composition should map to innovation outcomes (Table 1 of Section 2).

Our findings point to several generalizable insights. *First, we establish experimental evidence of the (co)existence of each of these sets of mechanisms acting alongside each other*. Our model of team composition explains almost a full 25% of the overall variation in innovation performance, suggesting that attempts to describe, predict, or prescribe optimal team composition should appeal to an integrative view of these mechanisms.

Second, our findings reveal important tradeoffs in the mechanisms driving the diversity-innovation link. Rather than these mechanisms necessarily being additive or complementary, emphasizing one mechanism tends to systematically impinge upon and downplay the others. For example, here we find that Social Processes dominated, with subgroups and faultlines (Lau and Murnighan 2005; Section 2.3) accounting for 54% of the variation in innovation outcomes explained by our model (Section 5.3). This importance of subgroups unavoidably worked against creating a broad span of non-redundant knowledge in these 4-person teams (Lazear 1999; Page 2019; Hamilton, Nickerson, and Owan 2012) as called for in theories of knowledge-as-input. So important were subgroups, having just a single team member from a specific field of knowledge had no statistical effect; statistical effects from different bodies of knowledge could only be detected when there were at least two from a given field. (Several reasons explain why Social Processes dominated in this context, see Section 6 for discussion).

Consistent with the Division of Specialized Labor, we found that expertise and training directly relevant to the problem (Business Computer Science, Design, and/or Engineering) were crucial to innovation performance. Specifically, only subgroups formed with these types produced the aforementioned large effects. This created a situation in which it would be inherently costly to allocate positions within a 4-person teams to those with intellectually distant and unrelated knowledge with the intent of fostering atypical knowledge combinations. Indeed, adding a team member with knowledge from fields that are somewhat intellectually distant from the innovation problem at hand (Health and Nursing, Humanities, Law, and Natural Sciences) did not have any direct benefit on innovation performance and only had an effect through contributing to overall group-level diversity. However, in this context, diversity per se and Diverse Recombination only accounted for about 6% of explained variation in innovation outcomes. Thus, each set of mechanisms emphasized distinct innovation priorities and each set of mechanisms somewhat crowded out the others (Section 6 provides a discussion).

The most striking specific finding follows the Social Processes perspective: "Balanced" teams with two members from one directly relevant knowledge domain and two from another outperformed all other teams. The performance benefits of balanced teams were more than twice the magnitude of having just one set of two-of-a-kind, indicating that the balance and faultline between subgroups accentuated benefits. Consistent with our interpretation of Social Processes, the evidence indicates that balanced teams (and subgroups more generally) led to overall better-functioning teams, more successfully coordinating and establishing momentum within this innovation sprint, generating both higher quality and novelty (Section 5.3). Also consistent with this interpretation, we find similar effects for another form of social category: we find analogous experimental evidence that gender-balanced teams outperformed those dominated by either men or women in this context.[3] Thus, the evidence is consistent with subgroups aiding in rapid sense-making and superior information processing, as predicted by theory (Section 2.3).

Therefore, apart from providing a novel integration and exploration of the three theoretical perspectives on diverse knowledge in innovation, the findings in this context particularly contribute to and reinforce the importance of theory on faultlines and subgroups (e.g., Gibson and Vermeulen 2003; Lau and Murnighan 2005; Carton and Cummings 2012). The prior literature has particularly measured demographic characteristics as the basis of identifying subgroups and faultiness (Meyer et al. 2014). Here we show that these theories and sociological mechanisms crucially apply to differences in knowledge, training, and expertise, as well. Of course, the social embeddedness of knowledge and expertise embodied in humans—shaped by years of training, selection, and socialization within fields—should not be surprising. Nor should the inherent challenges of integrating diverse knowledge sources be surprising (Lingo and O'Mahony 2010). However, the experimental evidence here demonstrates this social embeddedness of knowledge can even dominate the role of knowledge-as-input. Further, whereas prior literature on subgroups and faultlines has long debated positive versus negative effects, here we present clear causal evidence of (highly) positive effects in the case of subgroups and faultlines within a contemporary innovation sprint context in which novel products was conceived and designed. This research also aims to most centrally contribute to the small, but growing stream of research mapping how existing knowledge shapes the production of new knowledge and innovation (Fleming 2001; Lee Fleming and Sorenson 2001; Teodoridis 2018; Vakili and Kaplan 2021).

The findings here—and particularly the coexistence of multiple distinct *sets* of mechanisms and sharp tradeoffs among them—clarify that innovation managers need to explicitly consider their innovation

[3] These experimental results are consistent with recently published correlational evidence of gender in science (Yang et al. 2022).

priorities when designing teams. The finding of an especially pronounced role of the social embeddedness of knowledge underlines potential challenges and limitations of team inter-disciplinary problem-solving. Consequently, continued research on this topic could be informative in demarcating the comparative advantages of machine-based artificial intelligence in innovation (Korteling et al. 2021; Cockburn, Henderson, and Stern 2018; Brynjolfsson, Rock, and Tambe 2019)—particularly when compared to *teams* of humans.

2 Literature Review & Theory: Knowledge Diversity & Innovation

How does diversity of knowledge (training, expertise) affect team innovation performance?[4] This section reviews the multiple literatures relevant to this question and discerns three distinct theoretical perspectives on this question.

2.1 Boundedly Rational Innovative Search & the "Diverse Recombination" Perspective

Research in the economics and management of innovation has long characterized knowledge as an input to the innovation process, wherein knowledge inputs are feedstock to be transformed and synthesized into new knowledge and useful solutions (Xiao, Makhija, and Karim 2022). These processes are now widely referred to as knowledge recombination (Gruber, Harhoff, and Hoisl 2013) or recombinant innovation (Schumpeter 1942). This characterization of the relationship between knowledge inputs and knowledge outputs in the innovation process lends itself to a prediction that differences or diversity in innovation team members' knowledge can lead to the synthesis of potentially more creative, more novel solutions—and potentially high-quality breakthroughs (e.g., Fleming 2001; Kavadias and Sommer 2009; Uzzi et al. 2013).

Therefore, the Diverse Recombination perspective emphasizes the bounded cognition of individual innovators and the inherent uncertainty in the innovation search process. This characterization assumes that innovation is constrained by local search, where solution approaches are bounded by the knowledge a team begins with (Fleming 2001). According to this view, a team's knowledge diversity defines the scope of its search for new solutions. Consequently, diverse, and atypical knowledge combinations are expected to broaden the search, leading to more novel and potentially high-quality innovations.

In building models, theorists have sometimes assumed that knowledge combinatorics could yield increasing returns. For example, distinct knowledge elements A, B, and C might recombine into new solutions {A, B, C, AB, AC, BC, ABC}, which may further recombine on and on (e.g., Weitzman 1998).

[4] In contrast to many other activities studied in diversity literature, innovation uniquely involves tackling non-routine problems characterized by complex interdependencies (Fleming and Sorenson 2001; Herbert Simon 1991; Yayavaram and Ahuja 2008). These interdependencies often span diverse technological and knowledge domains (Leahey, Beckman, and Stanko 2017), making complex problem-solving shrouded in uncertainty. This uncertainty forces innovators to adopt search strategies to identify effective solutions (Abernathy and Rosenbloom 1969; Katila 2002). Unlike traditional organizational structures with fixed roles and schedules, which may be less effective, innovation teams require a dynamic approach. These teams often alternate between ad hoc interactions and isolated work, balancing exploration and exploitation to achieve their goals (Bruns 2013; Bunderson and Sutcliffe 2002; Cummings and Kiesler 2007; Faraj and Sproull 2000; Taylor and Greve 2006). Furthermore, the trajectory of innovative problem-solving is unpredictable, continuously evolving with emerging goals and possibilities (Cromwell, Amabile, and Harvey 2018).

In this perspective, knowledge inputs are the sum of non-overlapping knowledge quanta (Lazear 1999; Page 2019; Hamilton, Nickerson, and Owan 2012).

Building on this idea and theories of innovative search for solutions, in a seminal contribution, Fleming (2001) posits that common and routine knowledge pairings typically yield incremental advances, whereas rarer, more atypical combinations may prompt a wider search, resulting in lower average outcomes but with higher variability. This idea adds more formality to the intuition that atypical interdisciplinary knowledge combinations, for example, might sometimes yield breakthrough innovations, despite a higher risk of failure.

Fleming (2001) investigated these ideas in a wide cross-section of 17,264 patents, finding that more typical combinations of knowledge domains/categories are indeed correlated with more forward citations, indicating greater importance and impact. Conversely, he observed that the variance in forward citations was positively associated with less typical knowledge combinations. Uzzi et al. (2013) identified patterns aligning with Fleming's findings and found that only relatively minor degrees of atypicality were correlated with (i.e., statistically over-represented in) highly-cited scientific research. Their research indicates that among "hit" academic papers, there is a disproportionate number that blend conventional knowledge with a few unconventional or unrelated citations. Additional studies have importantly found similar correlations in other data sets, suggesting a link between atypicality and novelty and innovation (Hou, Li, and Lin 2021; Sternitzke 2009; Xiao, Makhija, and Karim 2022; Taylor and Greve 2006).

Other theories, however, raise questions about whether the breadth of knowledge is as crucial as depth for forging novel connections and achieving breakthroughs. For instance, theories of creativity that prioritize "depth over breadth" underscore the significance of extensive, domain-specific knowledge (Weisberg 1999). Kaplan and Vakili (2015) analyze 2,826 patents in the specialized field of nanotechnology, fullerenes. They use text analysis to show that the earliest patents introducing new topics tend to be associated with "local search" (greater depth in considering typical topic combinations). This view that depth rather than only breadth of knowledge might be important to finding novel solutions might also be read into Gruber, Harhoff, and Hoisl's (2013) finding that scientists (i.e., typically with narrower and more basic knowledge) tend to engage in more recombinant invention than do engineers. Vakili and Kaplan (2019) also raise the question of whether breakthroughs could come from teams with a good deal of knowledge overlap rather than breadth and differences (see Section 2.3) and they argue that the answer should depend on the context, and they indeed find that the correlations between team diversity and citations to patents differ across 4 different areas of technology patenting (MRI, RFID, stem cell, and nanotubes). Teodoridis, Bikard, and Vakili (2019) present evidence showing that the benefits of specialization over generality in individual mathematicians for producing academic breakthroughs vary with the pace of knowledge advancement.

2.2 A Burden of Knowledge and the "Division of Specialized Labor" Perspective

A distinct theoretical perspective from within innovation research views teams as vehicles for drawing together knowledge that aligns closely with the specific nature and structure of the problem being addressed. Instead of emphasizing atypical knowledge combinations, this perspective highlights the importance of applying deep specialist knowledge to the various aspects of a challenging innovation problem.

This view of innovation teams has been emphasized recently as part of the observation that as science and technology progress, it is increasingly difficult for lone inventors (the proverbial "renaissance man" (Jones 2009)) to have sufficient knowledge to address challenging innovation problems on their own (Jones 2009). Consistent with this idea, Jones (2009) finds that patents have increasing numbers of inventors and increasingly specialization of inventors in data between 1975 and 2001, across 414 major technological categories. He finds analogous differences in the cross-section of patenting categories when comparing fields with lower versus higher burdens of knowledge. Complementing these findings, Wuchty, Jones, and Uzzi (2007) examine 19.9 million academic research papers over 5 decades, revealing a shift from individual authors to team-based research. This team dominance is mirrored in their analysis of 2.1 million patents, representing a broad spectrum of technological areas.

The above explanation aligns too with lessons from the large and long-established literature on complex systems design, which emphasizes that challenging problems should be broken into manageable sub-problems or modules and tackled by specialist designers and problem-solvers (Alexander 1964; Parnas 1972; Simon 1991). Simon (1991) and others posit that organizing in a modular approach prevents the cognitive overload of any single individual by distributing complex decision-making across specialists. Moreover, partitioning the innovation task into subproblems improves focus and efficiency, enabling economies of specialization. These concepts now form a foundation for ideas in related research on Modularity (Baldwin, Clark, and Clark 2000) and Systems Engineering and Product Management (Marion and Meyer 2018; Simpson et al. 2006), in which the role of specialists is show or assumed to play a pivotal role in project success. However, here too, there is recent suggestion that the ways in which pre-existing knowledge enters into the production of new knowledge and new innovative solutions might rather nuanced. For example, Nagle and Teodoridis (2020) argue that teams without generalists may have difficulty including distant knowledge in their projects, enabling the sort of distant search described in recombinant diversity (Section 2.1).

Therefore, both this Specialized Division of Labor perspective and the earlier Diverse Recombination perspective consider knowledge as an input to the innovation process, and both assume limited cognition of problem-solvers. However, the Diverse Recombination perspective centers on seeking new solution methods and generating novel ideas through unconventional combinations. The Specialized Division of Labor stresses the challenge of imagining and executing high-quality innovations by aligning team members' expertise closely with the problem's specific nature and structure (to the extent these things are knowable ex-ante).

2.3 Subgroups, Faultlines, and the Team "Social Processes" Perspective

The third theoretical perspective has received greatest scholarly attention and theoretical development of the three (and consequently this subsection is somewhat longer than the earlier two). The starting point here is to focus on diversity in terms of differences in social categories and identity as a determinative characteristics shaping social dynamics (Roberson 2019a)—and thereby moderates team processes (e.g., Gibson and Vermeulen 2003; Mayo, Woolley, and Chow 2020; Pelled 1996; Pelled, Eisenhardt, and Xin 1999; Reagans, Zuckerman, and McEvily 2004). Past research has particularly tied social categories to demographic traits and functional or task-related attributes as a basis for social categories and identity (Bell et al. 2011; Bunderson and Sutcliffe 2002; Horwitz and Horwitz 2007; Van Dijk, Van Engen, and Van Knippenberg 2012). However, inasmuch as these differences apply to knowledge differences, then knowledge differences might shape and moderate the innovation process, apart from just acting as an input to that process (Figure 1).

<Figure 1>

i. Knowledge and Social Categories and Identities

Theories in the Social Process perspective hinge on differences in social groups manifesting in observable personal characteristics that are salient to social categorization (Pelled, Eisenhardt, and Xin 1999; Lau and Murnighan 2005). There are several reasons to suppose that differences in knowledge, training, and expertise should be both salient and observable markers of social categories. Knowledge often emerges over years of training, as in higher education, apprenticeship, practice, and professional association (Boone, Van Olffen, and Roijakkers 2004; Fouarge, Kriechel, and Dohmen 2014; Wiswall and Zafar 2021). Processes of sorting and socialization even often begin early in life within primary and secondary education and later with the choice of college major (Altonji, Blom, and Meghir 2012; Elman and Angela 2007). Formal training and socialization are also associated with inculcation of paradigms, epistemology, values, jargon, and methods (Pascarella and Terenzini 2005). Such long run processes in a field might then also result in a common sense of identity, status, language or jargon, and perspective (Hill et al. 2016; Roksa and Levey 2010).

ii. Differences, Faultlines, and Subgroups

Following these ideas, one perhaps simplest possible outcome of diversity might be a tension between "creativity-versus-coordination" as diversity increases. In this formulation, whatever benefits arise from diversity, greater differences can—roughly speaking—lead to frictions in coordination or communication (e.g., Huckman, Staats, and Upton 2009). A large number of plausible forms of frictions have been theorized in the literature, including: greater misunderstandings, lack of common language, clashing interpretations (Lix et al. 2022); lower social ties and social contact and integration (Dahlin, Weingart, and Hinds 2005); lower trust and cohesion (De Jong et al. 2021); reduced goal alignment and willingness to cooperate (Manata et al. 2021); greater social comparison and jealousy (Wang et al. 2016); negative emotions (Garcia-Prieto, Bellard, and Schneider 2003); clashing interpretations (Hoever et al. 2012); relationship conflict and dissent (Humphrey et al. 2017); ambient cultural disharmony (Paulus, Van Der Zee, and Kenworthy 2016); emergent social entrainment failure (Mayo 2022); and more.

A distinct stream of research instead emphasizes intra-team faultlines and subgroups. In doing so, this work predicts that the most intensive effects of diversity occur at intermediate levels of diversity, rather than most extreme levels of diversity (Lau and Murnighan 1998). Faultlines are defined as social distinctions that divide a team into subgroups based on one or more observable attributes salient to the group and individual identity (Jehn, Bezrukova, and Thatcher 2008; Lau and Murnighan 1998). The existing research especially emphasizes subgroups based on the team members' demographic alignment along one or multiple characteristics (Thatcher and Patel 2011, 1119). However, following earlier arguments, we might also expect differences in knowledge, training, and expertise to also be associated with social identification and self-categorization processes.

Theory in this area often points to the negative effects of faultlines and subgroups (Adair, Liang, and Hideg 2017). For example, Lau and Murnighan's (1998; 2005) seminal work first proposed a theory of faultlines as analogous to topological faults, which predicted intergroup conflict, reduced communication, and greater social distance in groups. Other theories suggest intergroup bias (Chiu and Staples 2013); in-group and out-group dynamics, exacerbated by status distinctions (Meyer et al. 2015; Yilmaz and Peña 2014); leading to reduced trust, team cohesion, increased conflict between subgroups, and lesser collective team identification (Privman, Hiltz, and Wang 2013; Van Der Vegt and Bunderson 2005; Flache and Mäs 2008). Conflict and reduced group morale might also lead individuals to withdraw and become less engaged, reducing effort (Cummings, Zhou, and Oldham 1993). The presence of multiple subgroups might also intensify these effects (Thatcher, Jehn, and Zanutto 2003) and produce unproductive competition for resources (Murnighan and Lau 2017; Polzer, Mannix, and Neale 1998). Open communication within subgroups in opposition to other groups can also produce more extreme positions and less productive exchange (Bezrukova et al. 2009).

At the same time, several counter-arguments for positive effects have been raised—many noted in the same studies that theorize negative effects. For example, groups with strong demographic faultlines and subgroups could benefit from shorter sensemaking processes in early formation stages (Lau and Murnighan 1998). Faultlines and subgroups and associated formation of coalitions could correspondingly help structure and organize a group (Lau and Murnighan 1998; Carton and Cummings 2013). Subgroups might also act with closer cohesion and deliberation, possibly pool resources, and unite in a single voice (Murnighan and Brass 1991), while also acting as supportive cohorts and means of finding common ground (Carton and Cummings 2012). Thus, these social mechanisms could potentially lead to enhanced information processing (Gibson and Vermeulen 2003) and accelerated consensus (Mäs et al. 2013).

Across different studies, the existing empirical research finds instances of both positive and negative associations between team outcomes and the presence and strength of faultlines and subgroups (e.g., Gibson and Vermeulen 2003; Thatcher, Jehn, and Zanutto 2003; 2003).[5] Distinctions in knowledge, training, and expertise, to our knowledge, have yet to be empirically probed from the lens of subgroups.

2.4 Summary & Research Questions

This section began with the question: How does diversity of knowledge, training, and expertise affect team innovation performance? The extant research has made substantial progress in proposing alternative theories on whether diversity might be productive or counterproductive. In the discussion, we discerned three distinct theoretical perspectives, each emphasizing different mechanisms. We summarize key predictions of these theoretical perspectives in Table 1, below.

<Table 1 Summary of 3 Theoretical Perspectives and their Alternate Predictions>

Although each theoretical perspective has its own lingering questions (see Sections 5.1, 5.2, and 5.3), our main thrust here will be to better integrate our understanding of each set of mechanisms in explaining innovation performance. This begins with making progress in establishing the existence, relative importance, and workings of the mechanisms shaping the knowledge diversity-innovation performance. Our main empirical strategy in the following sections will be to leverage distinct predictions, summarized in Table 1, as a starting point in our exploration of relevant patterns.

Beyond the key predictions, there are several larger questions lurking around our main points of enquiry. For example, first, prior work has often been couched in terms of whether diversity is productive or unproductive—or has a positive or negative effect. However, the foregoing theoretical perspective each suggests that innovation performance should vary with the degree or extent and type of diversity in question.

Second, none of the three theoretical perspectives a priori rules out the other perspectives. Therefore, on the basis of the existing theory and results, we predict that the mechanisms described in each of these theoretical perspectives should, in fact, coexist alongside the others. We expect that simultaneously considering and integrating these perspectives in our analysis might, therefore, add significantly greater power in describing and predicting the effects of knowledge diversity.

A third larger question to highlight is that each of these three theoretical perspectives points to rather distinct priorities in ameliorating and improving innovation. Further, each prioritizes rather different forms of knowledge composition of teams (Table 1). It is unclear, therefore, whether we should expect these mechanisms to be additive or even coherent with one another. Therefore, we also wish to better understand how each of these sets of mechanisms may impinge upon and relate to each other.

[5] Having found ambiguous results, numerous empirical studies and accompanying theoretical interpretations have begun to emphasize the possible complexity of effects, using moderating interactions, and alternative measures (e.g., Adair, Liang, and Hideg 2017; Schölmerich, Schermuly, and Deller 2016; Kaur and Ren 2023; Rico et al. 2012)

3 Experimental Research Design

The remaining sections empirically investigate the existence, relative importance, and workings of mechanisms shaping the knowledge diversity-innovation link (Section 2.4). The first and most challenging aspect of an ideal research design is to assemble a large group of individuals with widely diverse knowledge and training while this diverse knowledge still being plausibly applicable to the innovation problem. Ideally, individuals would be randomly assigned to comparable teams, working under controlled conditions, employing similar tools, and with the same innovation objectives. There would be a sufficient number of teams (observations) and sufficiently discerning measures of knowledge and objective innovation performance to support relevant inferences.[6] Ideally, the context would incorporate realistic features of the work task, work environment, and the workers themselves, to aid in supporting external validity. The promising advances in experimental and causal studies on diversity have not yet carried out such an exercise or squarely addressed our questions here (see Appendix A).

The gist of our approach is to randomly assign 872 subjects from a broad range of fields (Business, Computer Science, Design, Engineering, Health and Nursing, Humanities, Law, Sciences) into 218 teams of four. Each team tackled a novel IT product development challenge, creating a "next-generation" game that incorporates both virtual and physical interactions, over a three-week innovation sprint. Teams collaborated using a work-from-anywhere product development platform, where they developed a use case, business case, and technical architecture. The innovation outcomes, in terms of overall quality and novelty, were assessed by a large panel of C-level executives. This section elaborates further details.

3.1 Research Context

i. Work-From-Anywhere Product Development Platform

This research was carried out in partnership with an online collaborative product development platform hosted at a large top-40-ranked R1 university in the United States. The platform invites both alumni and students from all programs and disciplines to form teams and create new products utilizing contemporary information technology, which apart from usual personal computing, mobile, cloud, software and networking, also includes data science, embedded computing, and interactions with the physical environment, machinery—a technology stack often referred to as the Internet-of-Things (IoT). IoT remains in a nascent stage and many industry observers suggest there could be a plethora of yet-to-be-conceived products and services that foster connectivity among machines, infrastructure, consumer products, and more, leveraging the intelligence amassed from networked data collection (Patel et al., 2017).

By providing a wide cross-section of participants with opportunities to design IoT products ("applications"), the intent was to provide a greater understanding of changes in information technology and associated economic opportunities and to learn how these technologies may relate to participants' own

[6] The measurement and inference strategy should also account for the possibility that knowledge will be correlated with other traits (Section 5.2).

work organizations, while possibly kindling entrepreneurial opportunities. The program also provided experience in standard product development work practices from "use case" design to technical design and business case design. These goals were achieved through a series of part-time learning-by-doing product development challenges in which individuals assembled to work in teams competing with other teams. While people sign up as individuals, they are assigned to teams by the platform.

To make the program as widely accessible as possible, all activities were hosted on a work-from-anywhere online collaboration platform. The platform serves as a means of coordinating and organizing individuals into teams while also providing clear steps for proceeding with innovation challenges. The platform provided means of communicating and simultaneously observing and editing the same work outputs on all team members' screens. Asynchronous contributions were also possible. The platform also served as a work environment that provides steps and process, resources, and tools to carry out the work.

The platform was designed with the philosophy of "low floors and high ceilings" (Boaler, 2016), supporting a broad spectrum of participants ranging from non-technical people who might not have previously worked on Internet-of-Things projects, all the way to graduate students with high familiarity with IoT. For example, the platform work environment offered a series of steps and prompts, crafted to enable even a novice participant to engage in a progressive, step-by-step, hands-on learning experience. Experience showed that even individuals from a variety of technical or non-technical backgrounds could effectively proceed through these steps, but working in teams accelerated this process. At the same time, the platform allows for and accommodates however sophisticated and detailed of plans that might be developed by more seasoned professionals.[7]

ii. "Next-Generation Games" Innovation Challenge

The experiment was embedded within a 3-week innovation challenge in which individuals work in teams of 4 to conceive of and design a new "next-generation game." Next-generation games in this event were designed as games using both digital or virtual interactions, along with physical environments and/or physical things, including machines and objects of any kind. (Pokémon Go is an early successful example of such a game.) The problem of designing next-generation games of this kind maps closely to the emergent capabilities of the IoT technology. In addition to usual hardware, operating systems, applications software, networking, cloud, geo-location, etc., IoT adds greater emphasis on sensors, low-power networking, distribution of processing power across a network, imaging, embedding of systems in physical spaces, augmented reality, and control of physical systems. A completed submission included a use case design, technical architecture, and business case.

The submitted designs were intended to be a thorough and detailed proposal that could be evaluated from the perspective of whether further investments were warranted to take the design to the level of a working prototype. Therefore, the use case design included all game mechanics and rules and a complete

[7] The platform had over 5,500 participants from all disciplines and majors, including participants from each US state and 5 continents (before the program was finally discontinued in 2020, in association with the COVID-19 pandemic).

description of the experience and workflow. The technology architecture included a high-level representation of key working elements across the technology stack, including hardware, software, networking and cloud, data and data analysis including AI, physical sensors or actuators, machines and things, and environment. A drag and drop tool graphical tool was provided to facilitate sketching a technical architecture. The business case was focused on both broadly characterizing the target customer and market opportunity, and specifically quantifying and motivating whether investment in a working prototype was warranted, in terms of the market opportunity, the expected cost of producing a prototype (resources were provided to estimate costs), and clarifying which questions would be answered by producing a prototype.

iii. Feasible Contributions from Different Knowledge Fields

Despite the inherently technical aspects of innovating a new game using an IoT technology stack, several aspects of this innovation challenge made it meaningfully accessible to those from technical and non-technical backgrounds alike.[8] First, the task of imagining, conceiving of, and designing a new game concept is one that is, to a large degree, non-technical. Further, the parts of each solution submission—use case, business case, and technology architecture—involved non-technical aspects. Further, while the technology architecture part of the submission accepted detailed technical design sketches, the minimum requirements for submission involved following guided instructions for responding to questions regarding and a high-level graphical representation of the key components of the technical architecture design. The sequence of steps that culminate in a completed design were designed to be used by neophytes and experts alike. In addition to careful design of the step-by-step learning-by-doing framework, the collaborative platform enables working and learning with others while completing steps.[9]

iv. Plausible Conditions for Each Set of Mechanisms

Given our goal of investigating the co-determination of innovation outcomes by each of three theorized sets of mechanisms, as detailed in Section 2, it is crucial that the research context is not only representative of contemporary innovation, but also be conducive to the predicted mechanisms.

For instance, the conditions here are plausibly conducive to Diverse Recombination. The challenge of ideating a new game should appeal to a broad spectrum of people, including those who have played board games, sports, video games, or engaged in any other form of leisure activity. Additionally, the innovation process is designed to be inclusive, being accessible to contributions from both technical and non-technical participants.

In relation to the Division of Specialist Labor, the innovation problem can be clearly broken down in ways that align with specific specialist fields. For example, use cases relate to Design, business cases to

[8] Before running the challenge, the platform and process were trialed with dozens of individuals from across all university majors and disciplines to verify even individuals could successfully complete a minimum submission, regardless of background.

[9] The guided sequence of questions of learning-by-doing steps was trialed with dozens of students and staff members from across all majors to ensure that even individuals could complete the minimum requirements of a submission.

Business, and technical architecture to fields like Computer Science and Engineering. While it is not certain that these mechanisms will play a significant role, the necessary foundational conditions are in place.

Regarding Social Processes, several factors indicate why different categories of knowledge might be important in social interactions. Firstly, the nature of the problem and its components might themselves be highly salient to the importance of specific types of knowledge, such as technical versus non-technical. Secondly, the social context of interacting with individuals from the university could make differences in fields also be particularly salient, as affiliations within the university are typically divided by discipline and field. Moreover, the initial way in which team members learn about each other is through viewing each other's LinkedIn profiles, which emphasizes the significance of educational background.

3.2 Protocol and Randomization

i. Attracting Participants

Among the more than five thousand members on the platform at the time, 875 individuals signed up to participate in this challenge. (872 were assigned to teams of 4 and are part of the data analysis.) These participants received invitations to participate in the challenge through direct emails in the three weeks preceding the challenge. An initial email was sent, followed by a single follow-up. Clicking on the email sent participants to the platform, where they simply needed to click a sign-up button after signing in. They received a confirmation and notice that they would receive an email on the eve of the challenge start date to explain how to proceed.

In this invitation, the challenge was described in generic terms, explaining that the challenge would involve a problem related to the Internet of Things and would involve games, and that it would be suitable for people from all backgrounds and majors. It was also shared that the challenge would take place on a work-from-anywhere platform over 3 weeks, and it would be possible to participate outside of work hours. It was also clarified that people would be later assigned to a team that would include members of the university community.

Several motivations and incentives were disclosed in this initial communication. It was disclosed at this stage that the total cash prize pool was $15,000. Participants would also be able to continue to pursue their ideas after the challenge ended, should they desire (any designs and outputs created during the challenge would not be publicly disclosed). The invitation also highlighted the opportunity to learn and gain exposure to new opportunities in a learning-by-doing collaborative context. It was also reinforced that the technologies involved in the challenge were widely expected to be important for future economic opportunities in most every industry.

Of the 875 individuals who sign-up, 53% were graduated alumni and 47% current students. The participants included 39% females. All schools and colleges within the university were represented. The average age of the participants was 30 years.

ii. Random Assignment to Teams & Announcing Rules

After the 3-week sign-up period, there was a one-day gap before the 3-week innovation challenge began. During this window, the participant list was frozen and individuals were randomly assigned to teams of 4. The one incomplete team of 3 is not reported in the analysis. At the end of this gap day, as the challenge began, the team assignments were revealed on the platform by showing each team member (including the user, him or herself) as a set of pictures, drawn from Linkedin profiles, with accompanying summary text drawn from Linkedin appearing next to the picture. Clicking on the picture or adjacent Linkedin icon would send the user to the associated Linkedin profile for a more complete description of academic and professional histories of each teammate. (A requirement for signing-up was to have a Linkedin profile.)

Also at the beginning of the contest, the specific rules were revealed. Upon logging onto the platform, the specific problem statement appeared, which was to design a next-generation game, which would be a digital game involving both virtual and physical spaces and/or things. Submissions would use the tools and prompts of the platform and would include use case, technical architecture, and business case components as specified in steps outlined on the platform. The communications center and basic workings of the platform were also explained.

Greater precision in the structure of payoffs was also provided. In the provision of rules, participants also learned that the $15,000 prize pool would be shared equally by the top-3 teams. The next top-7 solutions received a runner certificate acknowledging their high-quality efforts within this Next-Generation Games challenge, signed and presented by the dean of the business school at the university. Participants also learned that final rankings and winners would be evaluated in double-blind scoring by a large panel of C-level executives recruited to the platform. Apart from numerical scoring, judges could also leave feedback. Team members in 4 randomly selected teams would receive tickets for a laser tag chain. Teams were also told that their ideas and designs would not be disclosed, in the event they should want to preserve secrecy were they to pursue the idea further after the challenge.

v. Three-Week Innovation Process

The team then proceeded for 3 weeks to complete their innovation designs. Work was carried out on a collaborative product development collaboration platform. A "communications center" always appeared on the right-hand side of the platform screen. Clicking on this tab enables a chat facility that allows 4-way chat messaging. In addition, the communication center has clickable icons for Skype and Google Chat, to immediately initiate communications with teammates.

The work itself was performed on the platform, where each of the members could simultaneously work on the design and all parts of the platform (akin to google docs, where multiple authors have their own pointers and can add to a document simultaneously). The platform's workspace is divided into the three parts of each submission: "Use Case," "Technical Case," and "Business Case" (Appendix B). Each section proceeds as a sequence of worksheets made up of guiding questions and design tools to facilitate the development process. (See Appendix B for screen shots.) The responses to questions and output of design tools compile into the full design. The Use Case section was designed for teams to provide a brief overview of the core idea, the target user, and the value proposition and to detail the gameplay and user experience. In the Technical Architecture section, teams provided details of the technological components of the design (sensors, actuators, analytics, networking, devices, and so on), first walking through text questions and then proceeding to a visual diagram using a drag-and-drop design tool. In the Business Case, teams addressed questions about the benefits, costs, and risks of building a prototype, approximations of the potential market, and pricing.

vi. Scoring & Evaluation of Designs

After the 3-week process, the structured responses, drag and drop design output, and any additional uploaded files were auto-compiled into integrated proposal documents that could be reviewed by judges in a double-blind process. To evaluate and score the designs, the university recruited several hundred C-level and VP-level executives to evaluate proposals via a platform-based interface. Each proposal was randomly assigned 12 evaluators. Each evaluator was randomly assigned 10 projects to evaluate. Evaluators logged onto the platform to complete evaluations online and was presented a dashboard list of each of their assigned projects to evaluate. The order of this lists presented to each evaluator was also randomized.

Each evaluator was asked to evaluate assigned projects in terms of overall quality and commercial potential. Evaluators were told to read each submission essentially as a proposal and sketch of an idea, and ultimately the goal would be to assess which designs deserved further investigation, evaluation, investment, and possible prototyping and testing. The side tab where judges read proposals on the platform presented a clickable 10-point scale. Evaluators likewise assigned a 10-point score for the novelty of concepts in the proposal. Rank order of submissions and winners was determined on the basis of average overall quality scores. These 10-point scores for overall innovation quality and novelty are the main dependent variables in the analysis to follow.

3.3 Data, Variables, and Descriptive Statistics

The data set analyzed in this paper includes 872 individuals organized into 218 teams of four. It was constructed by matching several data sources: data from the platform, data from university on education and training and fields of study, supplemented with educational data from Linkedin profiles, and inferences of demographic characteristics from various online analytics services. These data were meticulously matched and assembled at the individual level, from which we constructed team-level variables for our analysis of the 218 teams. Subsequently, this team-level data was matched with data on innovation outcomes and performance scores from the platform's back-end database. Descriptive statistics of main variables used in the analysis are detailed in Table 2.

<Table 2>

i. Measuring Innovation Performance

As described in Section 3.2.vi, each team's submission was evaluated by 12 C-level executives on a 10-point scale for overall quality and novelty. The overall innovation quality measure we use is the average of these 10-point score evaluations, *Quality*. We likewise use the average of 10-point novelty scores evaluations as our *Novelty* measure. Unsubmitted or incomplete proposals were assigned a score of zero.

ii. Measuring Knowledge and Diversity

To construct measures of fields of training, the university administration provided records of the most recent degree and year of graduation for alumni, and the undergraduate major for current students. We enhanced these records with LinkedIn data to complete the education histories. The dataset represented eight distinct fields of study: Business, Computer Science, Design and Media, Engineering, Health and Nursing, Humanities, Law, and Natural Sciences. We supplemented these data with LinkedIn profiles to capture cases where individuals received training from other institutions.

Previous studies have mapped diversity (an inherently multi-dimensional object) into some aggregate metric, such as the Blau index as in Section 4 (Meyer et al. 2014; Joshi and Roh 2009; Roberson 2019a; Harrison and Klein 2007). Mapping diversity into an aggregate summary of degrees of diversity can be especially helpful where there would otherwise be limited data and having a single metric consumes fewe degrees of freedom in one's statistical model. Here, we have a comparatively large data set (teams), allowing us to measure and map knowledge in a multitude of ways to investigate mechanisms. See details in Section 5.

iii. Control Variables

We construct a number of control variables related to demographic characteristics—race, gender, and age—as these have the potential to be correlated with knowledge and training (see discussion in Section 5.1). To discern gender and race from names, we used a gender-determining algorithm available via an online application programming interface (https://gender-api.com/en/about-us). This algorithm probabilistically determines gender based on first names using a large database of names and genders. Analogously, to classify race, we used a comparable ethnicity-determining algorithm, Ethniccolr (https://github.com/appeler/ethnicolr), a machine learning-based GitHub tool that maps first and last names to one of four race categories – "Asia Pacific Islander," "Hispanic," "Black," and "White."[10] The age of individuals was approximated by assuming students enroll in college at the age of 18.

[10] We hired two research assistants to verify, to the best of their ability, that the name-gender and name-race output of these algorithms appeared to be consistent with photos from publicly-posted LinkedIn profiles.

4 Overall Knowledge Diversity & Innovation

Before we proceed to our more detailed main analysis (Section 5), in this section, we report the overall relationship between experimental variation in knowledge diversity levels and innovation outcomes. The main findings here are: (i) intermediate diversity levels produce the highest innovation, (ii) well-functioning teams of intermediate diversity that generate the highest quality also generate relatively novel innovations, and (iii) if we more closely distinguish novelty from quality, we find that greater diversity leads to greater novelty when controlling for quality.

Panel I of Figure 2 reports the relationship between overall innovation *Quality* and variation in an aggregate summary measure of diversity, the Blau index. The Blau index is calculated as $1-\sum_{x=1}^{K} p_x^2$, where p_x represents the proportion of individuals in each designated knowledge category, such as Business, Computer Science, etc.. Figure 2 reports both a quadratic OLS estimate of the relationship along with a flexible, non-parametric estimate. The non-parametric estimate uses a locally-weighted linear model with an Epanechnikov weighting kernel estimated by a procedure described by Robinson (1988). Both estimates show the relationship peaks at intermediate diversity levels. The effect is quite large, showing about a 1-point conditional mean difference (on the 10-point scale) between teams with intermediate diversity versus entirely homogenous or highly heterogeneous teams.

The relationship between diversity and *Novelty* similarly peaks at intermediate diversity levels. Therefore, teams with intermediate diversity levels generate both greater *Quality* and greater *Novelty*, on average, relative to either highly heterogeneous or highly homogenous teams. To discern any separate effects on *Novelty*, Panel III of Figure 2 reports the relationship controlling for *Quality*: diversity leads to higher *Novelty*, when controlling for the *Quality* level.

<Figure 2>

These estimates reveal that intermediate knowledge diversity causes higher innovation in this contemporary product development context. However, they don't fully explain the mechanisms. Each of the theoretical perspectives (Section 2) might be used to explain the non-monotonic ("inverted-U") relationship, particularly if we consider that especially high levels of diversity can create frictions to weigh against whatever benefits of diversity. The upcoming analysis uses our wide variation in knowledge types, relatively high number of observations, and precise measurements—relative to what has been previously feasibly observed—to explore mechanisms.

5 Specific Levels and Types of Knowledge Diversity & Innovation

We now turn to our more detailed main analysis to make steps towards better understanding factors shaping the diversity-innovation link. (Section 6 discusses how results relate to theories of Section 2.)

5.1 Analysis Framework

The regression framework we employ distinguishes between two effects knowledge composition on innovation performance: the "direct effect" (of adding a piece of knowledge) and the "knowledge combination effect" (of how adding a piece of knowledge in the context of other types of knowledge). For example, there might be a direct effect of adding, say, an engineer to a team—and an added effects of having the engineer work alongside, say, a computer scientist. It is useful to distinguish these effects ot better explore and interpret mechanisms. Therefore, we use the following econometric framework:

$$y_i = \propto + \underbrace{\beta \cdot \mathbf{KnowledgeType}_i}_{\text{"Direct Effect"}} + \underbrace{g(\mathbf{KnowledgeType}_i)}_{\text{"Knowledge Combination Effect"}} + \theta_i + \varepsilon_i \qquad (1)$$

where y_i is the innovation performance outcome of team *i*. The $\propto$ term is a constant to be estimated. $\mathbf{KnowledgeType}_i$ is a vector representing numbers of team members of each knowledge type (Business, Design, Computer Science, Engineering, Health & Nursing, Humanities, Law, Natural Sciences). β is an 8-dimensional coefficient vector to be estimated. The term, $\beta \cdot \mathbf{KnowledgeType}_i$, is therefore the direct effect term. The function $g(-)$ operates on the entire vector of **KnowledgeType**, capturing the knowledge combination effect. We test this function using many specifications in the analysis to follow.

Beyond knowledge factors, other determinants are held constant largely as experimental controls, including: the innovation problem statement, time period, template and process, tools, communication platform. However, demographic and personal traits other than knowledge might also shape innovation, captured by the θ term. There could also be numerous zero-mean randomly distributed differences, ε.

5.2 Direct Effects of Adding Different Types of Knowledge

Here we estimate the direct effect (Section 5.1) of adding different individuals with different knowledge types to a team. We find here that only individuals from fields directly related to the problem have a direct effect on performance and the effect is non-linear, depending on the number of people added.

i. Innovation Quality

Results are presented in Table 3. All models are estimated using OLS with robust standard errors. Model (1) regresses overall innovation *Quality* on counts from each knowledge field (**KnowledgeType** in Equation (1)), also controlling for counts of team members with graduate degrees, and a constant. The coefficients on Business, Computer Science, Design, and Engineering are each positive and statistically significant (Section 3.3). Coefficients for Health & Nursing, Humanities, Law, Sciences, and Graduate Degree do not significantly differ from zero. The coefficient on *Grad Degree* is not significant. (If we break up graduate degrees by program and school, there is a weakly positive effect for graduate business (MBA).)

Of note, each significant knowledge type directly relates to the innovation problem and components of the submission (Section 3.2). These include the use case (Design), business case (Business), and technical architecture (Computer Science and Engineering). Unrelated, intellectually distant fields—Health and nursing, Humanities, Law, and Sciences—are each not significant. Thus, we adopt the terms "related" and "unrelated" in denoting the two sets. The importance of related fields most closely corresponds with the emphasis of related specialist knowledge predicted by the Division of Specialized Labor perspective.

<Table 3>

ii. Robustness

It remains possible that knowledge could be correlated with demographic characteristics, creating the possibility of biased estimates (see the θ term in Equation (1)).[11] To examine this possibility, we re-estimate our model with demographic controls (Section 3.4). In models (2) through (4), we alternately introduce our controls for gender (dummy), age (mean and range), and race (dummies). Model (5) then estimates the fully saturated model. In each instance, our main coefficients of interest on knowledge variables are unaffected. Results were also robust to our including industry dummies for work experience, based on LinkedIn industry codes—further affirming the relevance of our operationalization of knowledge.

iii. Innovation Novelty

We now investigate the effects on *Novelty*. Model (6) presents estimates without demographic controls, while model (7) includes them. In both cases, the coefficients are estimated to be similar to those in the preceding regressions with *Quality*. This result is consistent with well-functioning teams generating both high quality and high novelty relative to less well-functioning teams in this context. To better discern differences between the production of novelty and quality, model (8) re-estimates model (7), controlling for *Quality*. These estimates clarify that beyond whatever processes jointly generate both quality and novelty, adding individuals with Design training or graduate training has an additional effect on *Novelty*.

iv. Non-Linear Effects

Earlier estimates summarized the *average* marginal direct effects of adding one more team member of a given type. Here, we explore whether the effects depend on whether 1, 2, 3, or 4 individuals of the same type are added. Therefore, we replaced linear count variables with dummies corresponding to 1, 2, 3, or 4 members for each related field (i.e., 4 dummies for each of the 4 related fields). We include linear controls for unrelated fields, demographic factors, and the constant term.[12]

Coefficient estimates for each number for each related field estimated in this model are reported graphically in Figure 3. Each panel reports coefficients for the different related fields. Remarkably, in *each* case, adding just one person from a given related field has no effect; it is only when adding more than one that there is an effect. Having 2 or 3 of the same appears to generate the highest levels.

[11] This point is general to all studies related to personal characteristics (which cannot be randomly assigned), whether experiments or otherwise.

[12] Unrelated fields remain insignificant whether they are specified linearly or as a series of dummies, but linear controls consume fewer degrees of freedom.

Estimates with *Novelty* as the dependent variable produce similar results; again, factors leading to well-working teams lead to both greater quality and greater novelty. To detect any differences between *Novelty* and *Quality,* Figure 4 reports results for *Novelty* when controlling for *Quality*. Consistent with earlier results, adding those with Design training is significant. Again we observe there is only an effect when there is more than one.

<Figure 3>

<Figure 4>

5.3 Knowledge Combination Effects

Here, we investigate knowledge combination effects (Section 5.1) controlling for direct effects of adding different types of knowledge. After beginning by reporting several null results, our main finding here is that having subgroups—and especially balanced teams of 2-and-2—is critical to innovation performance. We also find evidence of incremental benefits from diversity per se.

i. Random Combinations & Innovation

As a baseline examination of possible knowledge combination effects, we examine how adding random combinations or draws from related fields of Business, Computer Science, Design, and Engineering (Section 5.2) affects innovation. We do so by constructing indicator variables for counts of whatever combination of 1, 2, 3, or 4 from these fields. As reported in Appendix C, coefficients on these variables are insignificant across multiple specifications, including controlling for direct counts of members from each field or not. Therefore, despite the positive direct effects of adding team members from these related fields (Section 5.2), random combinations are not effective. Any knowledge combination effects must be more specific than this.

ii. Complementarities and Interactions between Knowledge Fields and Innovation

We proceed to test the simplest form of knowledge combination effects: interactions or complementarities among fields. For example, following the Division of Specialized Labor perspective (Section 2.2), we might expect that adding multiple specialist team members could allow for a more productive division of labor, along with a greater span of applicable specialist knowledge. In Appendix C, we report results, adding interaction terms of related fields in our model, i.e., *Business* × *Design, Business* × *Engineering, Computer Science* × *Design, and Computer Science* × *Engineering.* We broadly find insignificant coefficients across multiple specifications. The lone exception is a weak positive interaction between Computer Science and Engineering.

iii. Main Finding: Subgroups and Innovation

A third sort of knowledge combination we consider is the particular knowledge configurations, in the sense of combinations of similar and different types (i.e., 2-of-a-kind, 3-of-a-kind, 4-of-a-kind). This may be relevant, for example, within the Social Processes perspective, where subgroups and the faultlines among them are predicted to shape team performance. To explicitly test the effect of different possible configurations and subgroups, we explicitly include in our model a set of dummies that reflect the exhaustive set of all possible knowledge configurations in 4-person teams.[13] We focus here on subgroups involving related knowledge.[14] (Configurations involving unrelated fields are insignificant, as will be shown in Figure 5.)

<Table 4>

For comparison purposes, model (1) in Table 4 begins by re-reporting the model of overall innovation *Quality* regressed on counts of different knowledge types (as in Section 5.1). Model (2) adds the series of dummies for knowledge configurations and subgroups. The coefficients on these dummy variables (except for fully homogenous 4-of-a-kind) are highly significant. The main finding is that having 2 or 3 of the same kind has (very) large effect on performance. However, the most striking of all is the effect of having a balanced team of 2-and-2, in which case the effect is larger. The estimated effects on *Quality* of subgroups from related knowledge fields are summarized graphically in Panel 1 of Figure 5. (Panel 3 of Figure 5 shows corresponding insignificant estimates of subgroups for related groups.)

<Figure 5>

Perhaps as important a finding, when controlling for knowledge configurations and subgroups from related fields, the coefficients on the variables controlling for counts from each field *each* become statistically insignificant. Remarkably, *this finding suggests that once configurations and subgroups are accounted for, the specific source of related knowledge is less important*. To further assess this point, model (3) drops controls for counts from related fields entirely and we find that doing barely affects the variation explained: the R^2 statistic in model (3) drops to 0.24, barely different from the R^2 statistic of 0.25 before dropping variables in model (2).[15] Therefore, controls for specific fields provide almost no explanatory power, once accounting for subgroups from the related fields (Business, Computer Science, Design, Engineering).

Models (4) through (6) estimate analogous models for *Novelty*. Given that well-working teams here tended to generate both quality and relative novelty, results are similar. In addition, consistent with earlier results (Section 5.2), coefficients related to design and graduate training are again both positive. To more closely examine possible differences between the generation of Novelty and Quality, we re-estimate the model for *Novelty* controlling for *Quality,* as in models (7) through (9). This highly stringent estimate of effects (other than whatever effects are shaping quality and novelty at once) detects positive effects on novelty of engineering and business training, apart from design and graduate training. Once *Quality* is controlled, there is no additional effect on Novelty from knowledge configuration and subgroup variables.

[13] For example, with 3-of-the-same-type configurations, there will necessarily only be one configuration of AAAB. However, in the case of two of the same, there can be AABB or AABC. There is also the possibility of all four the same, AAAA, or all four different ABCD.

[14] For example, there is one possible configuration of three-of-a kind (AAAB). Teams with two-of-a-kind can be or AABC, etc..

[15] Controlling for field using the series of dummies for specific numbers (1, 2, 3, or 4) for each field (as in Section 5.2) finds similar results as those reported here for linear controls of counts from different fields, however standard errors become quite large.

The estimated effects on *Quality* of subgroups from related knowledge fields are summarized graphically in Panel 1 of Figure 5. (Panels 2 and 4 graphically show insignificant estimates of subgroups, whether in relation to related or unrelated fields on Novelty, when controlling *Quality*.)

These results are consistent with knowledge configurations and subgroups simultaneously shaping overall quality and relative novelty of well-working teams. Once accounting for well-working processes (as by controlling for *Quality*), subgroups have no added effect on novelty. There appear to be other incremental sources of novelty beyond having well-working teams, such as when adding knowledge from Design, Engineering, Business, and graduate training.

Additional Test for Social Processes: Gender-Balanced Teams Also Outperform

The Social Processes perspective suggests that this pronounced effect of subgroups derives from distinct social categories and identities, rather than necessarily from knowledge per se. We already found evidence consistent with this point (i.e., once subgroups are accounted for, the specific source of related knowledge is less important). Further, If this interpretation is correct, we might expect analogous effects for other salient social categories, even if unrelated to knowledge. Here we test for such subgroup effects with gender.

Model (1) of Table 5 regresses innovation *Quality* on indicators for the presence of 1, 2, 3, or 4 women on a team. (The indicator for zero women is redundant due to the constant and is therefore excluded.) As reported in the model (1), we indeed find that gender-balanced teams outperform all other configurations and exceed all-men teams by over a full point (1.19, s.e. = 0.60). These estimates only become more significant when controlling for knowledge and other demographic controls, as in model (3), or controlling for the full set of knowledge configurations, as in model (4). Also consistent with prior results, gender-balanced teams also produce higher novelty, as in models (5). Also consistent with prior results, once we control for *Quality, there is no effect of gender* balance on *Novelty*, as in model (6).[16] Therefore, these results are consistent with subgroups better enabling these teams and that differences in social categories—either with regard to gender or with regard to knowledge—produce analogous effects.

<Table 5>

Additional Test for Social Processes: Subgroups Help Initiate Coordinated Action

The Social Processes perspective suggests that clear subgroups and faultlines can lead to benefits in rapid sense-making. Observable traits serve as cues around which members can form coalitions, begin to organize, and initiate communications (Section 2.3). If true, we should find evidence that teams with subgroups were better able to initiate some minimum level of coordinated action in this 3-week innovation

[16] Note that in model (6), as the model of *Novelty* conditional on *Quality* is estimated with still greater precision when accounting for gender subgroups, and the coefficient on homogenous knowledge (AAAA) becomes significantly negative. This result is consistent with earlier analysis suggesting that despite the benefits of subgroups, that 4-of-a-kind—complete homogeneity—is not conductive to innovation here.

sprint. To test this idea, we analyze whether the presence of subgroups affected the probability of simply mobilizing a team to make a successful submission and to achieve a score greater than zero.

For comparison purposes, model (1) restates coefficient estimates for earlier regressions of overall innovation *Quality* on subgroup variables. Model (2) replaces the dependent variable with an indicator for receiving a score greater than zero. We again see that subgroup variables are significant, forming a similar pattern of balanced teams of 2-and-2 being most likely to submit a proposal and receive positive points. There is no direct effect of specific types of knowledge (i.e., the counts from each knowledge type) on the probability of receiving a positive score. Although the model fit, in terms of Adjusted-R^2 is not quite as high as in model (1), it only drops from 0.17 to 0.12. Therefore, a large fraction of variation in model (1) comes from simply explaining the ability to mobilize a team to achieve more than zero score. These results are consistent with the Social Processes perspective.

For comparison, we also report regressions for *Quality* (model 3) conditional on successfully submitting and receiving a score greater than zero. It is more difficult to interpret these regressions, with just 74 observations (34 percent of teams successfully made completed submissions) and these are highly selected and necessarily shaped by endogenous sources of variation. Nevertheless, we see in regressions of Quality conditional on a completed submission, point estimates on each of the subgroup variables are positive and that on two from the same field is significant at $p = 0.10$ (model 3). These results are consistent with subgroups leading to relatively well-working teams, better able to prepare completed proposals and able to achieve incrementally higher quality. These results are consistent with the Social Processes perspective.

We proceed to use this smaller set of 74 endogenous observations to draw distinctions regarding *Novelty*, when we condition on making a completed submission in model (4). Again, to draw the clearest distinctions between Novelty and Quality, we also control for *Quality*.[17] Here we see consistently negative point estimates on subgroup variables: the larger the size of the subgroup, the larger the negative effect on *Novelty* (model 4). The negative coefficient on three of a kind from the same field, 0.87, is significant at $p = 0.05$. These results are consistent with the benefits of Social Processes in supporting successful team processes (and quality) coming at some incremental cost in novelty.

We continue to see a strong relationship between training in Design and graduate traing related to *Novelty* when conditioning on making a successful submission and controlling for Quality. This result continues to be consistent with these incremental effects from different types of knowledge on novelty being from a distinct set of processes than how subgroups contribute to well-working teams.

[17] Not controlling for Quality leads to statistically similar, but less precise estimates.

iv. Diversity per-se and Innovation

We proceed to test for whether diversity per se plays a role in shaping outcomes, as the Diverse Recombination perspective predicts. Earlier, in Section 4, we estimated the relationship between innovation outcomes and diversity per se, as measured by the Blau Index; however, here we do so in the context of our fully-specified econometric model—accounting for other effects. Results here are reported in Table 7.

Model (1) estimates the model of *Quality* on the Blau Index, controlling for direct counts of knowledge, knowledge configurations, and subgroups variables, and demographic controls. This model effectively replicates earlier analysis, while adding the Blau Index measure of diversity per se. We find the coefficient on the Blau Index is positive and sizeable, but statistically insignificant. Either diversity per se indeed has zero effect on innovation, once controlling for other mechanisms, or our test is underpowered. To evaluate this latter possibility, we re-estimate model (1), but drop the insignificant controls for counts from specific knowledge domains, as in model (2).[18] Doing so leads the point estimate on the Blau Index to be virtually unchanged, 3.49 in model (2) versus 3.68 in model (1), but the standard error becomes considerably smaller, 1.76 in model (2) versus 2.75 in model (2). The result is consistent with model (1) being statistically underpowered and diversity per se has an incremental positive effect on *Quality*.

<Table 7>

We perform analogous tests replacing the dependent variable with *Novelty* in model (3). We find a similar positive and significant coefficient on the Blau Index. Model (4) proceeds to more closely discern any differences between *Novelty* and *Quality*, by adding *Quality* as a control. We find that diversity per se leads to greater *Novelty*, when stringently controlling for the particular level of *Quality*, as the coefficient on the Blau index is positive and significant at 0.53 (s.e. = 0.26). These results are consistent with part of innovation outcomes explained by Diverse Recombination, with an effect on both *Novelty* and *Quality*. Importantly, this effect is estimated alongside the effect of earlier mechanisms.[19]

5.4 Explanatory Power and Magnitude of Theoretical Perspectives

Our model accounts for a large share of variation in innovation outcomes, based solely on measures of team composition. The models in Section 5.3.iv explain approximately a quarter of all variation (i.e., *R-squared* statistic = 0.249). Among the three perspectives of Section 2, the Diverse Recombination approach explains the least. As shown in Table 7 of Section 5.3.iv, removing the Blau Index from model (2) reduces the (unadjusted) *R-squared* statistic from 0.249 to 0.235, representing about 5.6 percent of the explained

[18] Recall that once controls for knowledge configurations and subgroups (from related fields) were included in our model, the controls for direct counts of those from each related field became insignificant (Section 5.3.iii). Further, controls for counts from unrelated fields are insignificant across all models.

[19] Research and seminal theory within the Diverse Recombination perspective (Section 2.1) also predicts there should be greater variation in quality of outcomes with greater diversity (L. Fleming 2001). We tested for this this by re-estimating models in this section while parameterizing the estimate of variance as a function of the Blau index, where conditional mean and this conditional variance were estimated simultaneously by maximum likelihood. Although point estimates of on the the Blau Index model of variance is found to be positive, it is statistically insignificant.

variation. By contrast, the demographic controls account for roughly 10.1 percent of the variation or about 40.5 percent of the explained variation.

If is more difficult to provide an unequivocal breakdown between the Division of Specialized Labor and Social Processes perspectives. This is because the significance of subgroups and faultlines (related to Social Processes) is limited to directly related fields (relevant to the Division of Specialized Labor). Thus, while the single most important result relates to subgroups, these two mechanisms are intertwined. Jointly, these two perspectives account for the remaining 53.9 percent of explained variation.

<Figure 6>

6 Summary & Discussion of Results

6.1 Summary of Results

The main results from the preceding analysis are summarized in the following table, while relating findings to predictions of the 3 theoretical perspectives (Section 2). Our analysis began by documenting that teams with intermediate diversity had the highest innovation performance (Section 4). A series of more precise mapping of experimental variation in knowledge to innovation outcomes in Section 5 allowed each of the mechanisms to be more directly characterized.

Table 8: Summary of Main Empirical Results & Consistency with Theoretical Perspectives

	Result	Diverse Recombination Perspective	Division of Specialized Labor Theory	Social Processes Perspective
1	Intermediate levels of overall diversity led to highest innovation performance; controlling for innovation quality, novelty increases with greater diversity (Section 4)	✓	✓	✓
2	Adding team members with knowledge that is most related to the problem (Business, Computer Science, Design, Engineering) boosts innovation performance; more intellectually distant knowledge (Education, Health and Nursing, Humanities, Law, Natural Sciences) has no such direct effect (Section 5.2)		✓	
3	Subgroups (of those with related knowledge) are critical to innovation performance; adding just one team member with related knowledge has no effect; subgroups are more important than the specific source of knowledge (Section.iii)			✓
4	"Balanced" teams with 2 of one related knowledge and 2 from another outperform all other configurations; the effect is more than double the effect of one subgroup of 2; similar effects are found in gender-balanced teams (Section 5.3.iii)			✓
5	The effect of subgroups on *Quality* largely (but not entirely) comes from teams with subgroups more effectively coordinating and mobilizing (successfully making a completed submission) (Section 5.3.iii)			✓
6	Although subgroups help *Novelty* from effective coordinating and mobilizing, having multiple members of same knowledge reduces Novelty all else being equal (Section 5.3.iii)	✓	✓	✓
7	Diversity per se leads to higher innovation *Quality* and even higher *Novelty* (Section 5.3.iv)	✓		

6.2 Evidence of Diverse Recombination—and Knowledge-as-Input to Innovation

The Diverse Recombination perspective (Section 2.1) predicts diversity and atypical knowledge combinations will lead to more novel discoveries, and possibly higher-quality breakthrough advances. Using our field experimental framework, we found evidence consistent with this mechanism working alongside others theorized in Section 2. Most importantly, we found that diversity per se, as summarized by the Blau index, once accounting for other effects, contributes to greater novelty of innovation outcomes and, to a slightly lesser degree, to greater overall innovation quality (Section 5.3.iii). Also consistent with this perspective, we found to the extent there are larger subgroups

Throughout the analysis, we also find a general tendency for homogeneity to work against the production of novelty, particularly when controlling for the quality. For example, in model (6) of Table 5, the larger the (homogenous) subgroup, the lower the novelty, when controlling for quality. More broadly, in Figure 2 of Section 4, we saw that greater diversity per se, as captured by the Blau index, is monotonically related to greater novelty when controlling for quality. Therefore, there are multiple patterns in the experiment of a contemporary product development innovation spring affirming the existence of Diverse Recombination (Section 2.1) mechanisms—acting alongside others.

The magnitude of effects was relatively small in this context; one estimate indicates that Diverse Recombination accounts for about 5.6 percent of explained variation (Section 5.4). The relatively small role of Diverse Recombination here, however, does not diminish the demonstration of (co)existence of this mechanism within our experimental framework. Further, the generation of novel, breakthrough innovations might also, itself, be a rare outcome.

To the extent the small effect of Diverse Recombination here runs counter to usual intuitions, several points might be recognized. Here, we focused on the application of technologies within a contemporary product innovation sprint, in which we sampled each project attempted. This contrasts with the prior in prior studies on studying relationships within published academic research or patented technical invention teams (Uzzi et al. 2013; Fleming 2001); so, we should expect important differences. Further, we might expect in this context, despite the accessibility of the problem to all types, those beginning with some measure of relevant knowledge appear to have had an important advantage. Thus, depth trumped breadth here in devising and designing new concepts (e.g., Kaplan and Vakili 2015).

6.3 Evidence of The Division of Specialized Labor—and Knowledge-as-Input to Innovation

The Division of Specialized Labor perspective (Section 2.2) predicts that optimal differences in team knowledge will map closely to the nature and decomposability of the problem. In our findings, we indeed found that adding team members with directly related knowledge—Design, Business, Computer Science, Engineering—proved critical to performance (Section 5.2). These related types of knowledge, most intellectually proximate to the problem at hand, also mapped closely to the decomposability of the problem (i.e., the separate components of a submission: use case, technical architecture, and business case). There were no such direct benefits we could detect from adding team members from more intellectually distant fields like Health and Nursing, Humanities, Law, and Natural Sciences. (Unrelated fields only contributed via their contribution to diversity per se and Diverse Recombination, rather than in and of themselves.)

Initially, the direct effect of adding relevant team members appeared substantial, accounting for a significant portion of variation in outcomes (Section 5.2). However, the explanatory power of adding team members with relevant specialist knowledge was eclipsed by the role of subgroups (Section 5.3.iii): Once accounting for subgroups, the type or source of knowledge provided little explanatory power (long as it was from one of the fields of related knowledge). Further, other predictions of the Division of Specialized Labor perspective were difficult to detect. For example, we found little evidence of complementarities or positive interactions from having multiple sorts of related knowledge (Section 5.3.ii).

The results, therefore, on the one hand, appear to strongly affirm the idea that relevant specialist knowledge that is closely aligned to the structure and decomposability of the problem plays a central role in determining the optimal knowledge composition of teams. On the other hand, this effect along with other effects related to knowledge-as-input (Section 6.1) exist but play a secondary role in this context. As we will argue below, this is the result of being somewhat crowded out by the dominance of Social Processes and the role of intra-team subgroups and faultlines.

6.4 Evidence of Social Processes—and Knowledge Differences Shaping Team Processes

The Social Processes perspective emphasizes that differences in social categories and identity can shape team processes. The relevance of differences (faultlines) versus similarities (subgroups), the traits and measures best capturing social categories, and whether effects are positive or negative remain hotly debated (Section 2.3).

The most important empirical finding of all in this contemporary product development context was of the dominance of Social Processes over other views in explaining variation in innovation outcomes. Whereas the earlier theoretical perspectives stressing knowledge-as-input to innovation processes focus on the input of non-redundant knowledge (to generate diversity and atypical combinations, or to span the relevant scope of a problem and to enable a division of labor), here we found that intra-group subgroups and faultlines play a crucial role in shaping outcomes. For example, having just one person from a given knowledge field has no statistically discernible effect on innovation performance (Section 5.2); it is only when there is at least 2 of the same kind. Again, once accounting for subgroups from whatever areas of relevant knowledge, the specific source of knowledge provides little explanatory power (Section 5.3.iii).

The specific knowledge subgroup configuration that generated the highest performance was a "balanced" team consisting of two members from one related field paired with two from another related field (Section 5.3.iii). Apart from subgroups, this result also indicated a role for faultlines, as the effect of two subgroups of 2 exceeded the effect of twice the magnitude of one subgroup of 2.

Just as in the theory, the findings are consistent with knowledge and training acting as a salient social category shaping team processes. First, as earlier mentioned, the specific type of knowledge are far less relevant once subgroups are accounted for. Second, patterns in relation to subgroups suggested that subgroups simply led to well-working teams, as we found similar effects on the generation of both quality and relative novelty—it was a general improvement in performance (distinct from other effects, which shaped quality and novelty separately). Third, when testing on another form of social category that is separate from knowledge and teams—gender-balanced teams—we found the same result (Section 5.3.iii). Analogous correlational evidence of benefits of gender balance have also been presented in recent research (Yang et al. 2022). Fourth, consistent with the theory, subgroups and balanced teams are better able to mobilize and initiate collective team action within the 3-week innovation sprint; much of the benefit of having intra-team subgroups was manifested in a greater probability of making a completed submission at all (Section 5.3.iii). This finding is consistent with the theory (Section 2.3), which emphasizes how subgroups and faultlines can aid in shorter sensemaking processes in early formation stages (Lau and Murnighan 1998); help structure and organize a group (Lau and Murnighan 1998; Carton and Cummings 2013); and help decision-making and information processing (J. K. Murnighan and Brass 1991; Gibson and Vermeulen 2003; Mäs et al. 2013). We do observe some countervailing costs of forming subgroups, such as the earlier-mentioned reduction of novelty, conditional on quality, with larger subgroups; and we might also imagine that any number of other negative effects conjectured in the theoretical literature might play some role. However, the overwhelming effect of mobilizing teams and allowing to achieve some level of coordination and joint production is the dominant effect.

There are several reasons why Social Processes might have played a dominant position in this context. For example, just as in many cases of deliberately assembling diverse teams, these teams are newly-formed and built for a specific purpose. The newness of teams will necessarily make sense-making and incipient organization especially important. Also, while the innovation exercise here is lengthy and challenging relative to what might appear in a lab experiment, this 3-week development process is analogous to a fast-paced sprint project with an add hoc assembled ideation and design team or first steps within a larger development project (Kelley 2001; Goh, Goodman, and Weingart 2013; Burke and Morley 2016). This newness and time pressure is especially relevant to a central idea within the Social Processes perspective—that subgroups accelerate sensemaking within teams. Theory in this area highlights how recognizable attributes for forming subgroups can accelerate sensemaking, thus facilitating organization, activation, and problem-solving (Section 2.3). Shared language, values, identity, perspectives, epistemology, and methods can accelerate productive mobilization, idea exchange, and raise the productivity of teams as information processing organs within an innovation sprint, as we observed here. In this context, where all participants graduated from the same university in distinct disciplines and where LinkedIn profiles prominently featured educational achievements, we should expect that disciplinary training would have likely naturally played an outsized role as a visible marker for organizing subgroups. This exercise also purposefully did not include a manager or other more deliberate mechanism to substitute for these Social Processes mechanisms. We should, of course, expect the balance of the importance of the three sets of mechanisms should vary by context.

6.5 Co-existence of Mechanisms & Sharp Tradeoffs

In providing field experimental evidence of the existence of each of these sets of mechanisms acting simultaneously alongside the others, the findings make a case for better integrating theory that has largely been developed across separate and independent literature. The predictive power of doing so appears to be quite high; our model of team composition explains roughly 25% of variation in both innovation quality and novelty here.

Moreover, failing to consider the multiple sets of mechanisms as potential co-determinants can make it difficult or misleading to evaluate the knowledge diversity-innovation link, and to extrapolate theoretical insights beyond a particular setting. For example, in Section 4, we were able to exploit experimental variation in diversity to show that teams with intermediate degrees or levels of diversity outperformed other teams. However, as we discussed, it is possible to interpret that fact as somehow consistent with each of the three perspectives if taken in isolation. Despite this being an experimentally developed finding, additional careful testing was required in Section 5 to better discern mechanisms.

It should also be especially important to continue to build theory on an integrated basis, as the results here indicate that attempts to promote one set of mechanisms will likely face sharp tradeoffs with other mechanisms. Although these mechanisms coexist, they are hardly simply additive or complementary. For example, enhancing Social Processes dominated, and the role of subgroups and faultlines (Lau and Murnighan 2005; Section 2.3) accounted for 54% of explained variation in innovation outcomes (Section 5.3). But choosing team members to boost subgroups will inevitably reduce the span and diversity of non-redundant knowledge (Lazear 1999; Page 2019; Hamilton, Nickerson, and Owan 2012) as emphasized by both theoretical perspectives emphasizing knowledge as an input. Here, consistent with the Division of Specialized Labor, we found that expertise and training directly relevant to the nature and decomposability of the problem (Business Computer Science, Design, and/or Engineering) were crucial to innovation success. However, by definition, teams populated with these types would be the "usual suspects" (Lakhani and Jeppesen 2007) and would necessarily emphasize depth over the breadth of knowledge (Kaplan and Vakili 2015) called for in Diverse Recombination. Therefore, there appear to be many inevitable and sharp tradeoffs.

6.6 Team Composition and Processes of Generating Quality and Novelty

Just as the theoretical perspectives stressed either knowledge-as-input to or knowledge-as-moderator of innovation processes, the results suggested two broad modes in which quality and novelty were generated.

For moderating factors of subgroups and faultlines, there was a close correspondence between effects on both the level of quality and relative novelty generated. This follows from well-working teams simply working better and generating more of the desired outcome in this context. (Although relative novelty might not objectively be desired in all contexts, most every concept in the case of next-generation games would have been novel at the time and a desired outcome.) Much of this close correspondence was simply from the fact that well-working teams were able to initiate coordination and team collective action to make successfully completed submissions (Section 5.3.iii), meaning non-zero scores on both quality and novelty. The close correspondence of novelty and quality outcomes was recorded throughout, consistent with the dominant of moderating effects of Social Proceses in this innovation sprint.

For other factors shaping outcomes, as related to knowledge-as-input, there is evidence of a divergence in how quality and how novelty are generated. For example, among teams who successfully made completed submissions, those well-working teams with subgroups tended to receive higher quality scores, while larger subgroups were associated with lower novelty (Section 5.4). Further, we found that having team members from Design or graduate training contributed to higher novelty; whereas quality was not directly affected by these factors (Sections 5.2, 5.3, and 5.4). By contrast, we noted that quality could be improved by the presence of an MBA on a team (Section 5.2). We also see that diversity per se had a greater effect on novelty than on quality (Sections 4 and 5.3.iv).

6.7 Potential Limits to the Assembly of Diverse Knowledge by Teams of Humans

As much as these results provide an integrative view for understanding the multiple mechanisms at work that link knowledge diversity and innovation, these finding also raise important questions regarding joint problem-solving and innovation by multiple people within teams. The three sets of mechanisms are somewhat antagonistic to one another. Further, the dominance of Social Processes here underlines the point that human knowledge will necessarily be socially embedded—the results here underline the extent and importance of this social embeddedness in shaping the scope and means by which knowledge be harnessed and processed by humans working within teams. These points both underline the need for continued research into the "grammar" by which humans can combine distinct knowledge sets – and also possible limits that humans will face when attempting to assemble diverse knowledge. Better understanding of these points is likely to become more pressing in order to understand whether, how, and when machines and AI may have comparative advantages in combining diverse knowledge to solve problems and innovate (Korteling et al. 2021; Cockburn, Henderson, and Stern 2018; Brynjolfsson, Rock, and Tambe 2019). For example, there is little track record of humans successfully engaging in inter-disciplinary research in any systematically successful way. These questions deserve careful study.

7 Conclusion

The pre-existing knowledge (training, expertise) assembled around a project is among the primary determinants of the innovation. Nonetheless, the logic and grammar by which knowledge shapes the innovation "production function" within teams remains relatively uncharted. By comparison, models of other determinants of innovation—such as incentives, governance, and organizational context—are, today, highly elaborated and precise in predictions. Further, despite considerable interest in the question of knowledge diversity, theory has proceeded in a fragmented fashion to this point and there is yet little discriminating causal evidence to adjudicate among competing predictions in the theoretical research. Thus, strictly speaking, it has not been possible to offer clear prescriptions to innovation managers forming teams (Vakili and Kaplan 2021).

We posited that mechanisms described in three theoretical perspectives—Diverse Recombination, the Special Division of Labor, and Social Processes—would co-determine innovation performance (Section 2.4). This is especially notable as it has bee largely separate literatures studying these distinct sets of mechanisms (Section 2). We argued that gaining a better demonstration and understanding of the (co-)existence of these mechanisms, their relative importance, their workings—and how they might relate, interact, or impinge on each other—would open a greater understanding of optimal knowledge diversity in teams.

We first reported experimental estimates of the causal relationships between knowledge diversity and both overall innovation quality and novelty in a contemporary product development 3-week innovation sprint (Section 4). These estimates are themselves a contribution to the literature, as it has been inherently challenging to assemble large numbers of widely diverse individuals from many fields, randomly assigned to teams working under controlled conditions for which measures of innovation performance are available. The experimental approach adopted in this study is a methodological advancement in the case of study diversity of knowledge, training, and expertise. The controlled yet realistic setting, combined with a diverse participant pool, provides a robust platform for discerning the causal effects of knowledge diversity. This approach might perhaps serve as an approach for future research in this field, particularly in examining complex interactions within team dynamics.

Our most basic general contribution was to present existence evidence of mechanisms associated with each of these views (Section 5), acting alongside the others, within a causal field experimental framework. By presenting experimental evidence that demonstrates the coexistence (Sections 5 and 6) of mechanisms, this study bridges gaps between previously isolated theoretical frameworks. Our model, based only on team characteristics and composition, could account for a roughly a quarter of all variation in innovation performance. Apart from these three sets of mechanisms coexisting and co-determining innovation performance, these mechanisms clearly impinge upon and delimit the action of each other, as was discussed in Section 6.5. In this context, Social Processes dominated (Sections 6.4 and 6.5). These are somewhat unforgiving tradeoffs that innovation managers need to reconcile with their problem and context. (Also see Vakili and Kaplan (2021) on related points). These results underscore the need for an integrated approach to understanding and leveraging knowledge diversity in innovation teams. This integration offers a more comprehensive understanding of the tradeoffs and synergies between different knowledge dynamics.

We found, more specifically, that balanced teams of two-and-two from related knowledge fields outperformed all other teams within this contemporary product development context organized as a 3-week innovation sprint. We showed how these results indicate that knowledge differences relate to mechanisms of Social Processes (Section 5.3.ii and Section 6.4). More deeply, these findings affirm the importance of the social embeddedness of human knowledge, and that training and expertise are closely associated with distinct social categories and personal identities. What is most striking here is the extent to which these points play a first-order role in shaping innovation, even more than knowledge-as-input. Among supporting tests to validate findings, we found analogous experimental effects of high performance in gender-balanced teams (Section 5.3.iii). Yang et al. (2022) have found similar patterns in naturally-occurring data in scientific research teams.

For practitioners, this research offers critical insights into team composition and management. The emphasis on social processes and balanced teams suggests that innovation managers should be mindful of not only the knowledge makeup of teams but also the social structure and dynamics these compositions engender. These points are, on the one hand, present in prior literature, albeit across largely separate literatures. However, the present study clarifies the extent and sharpness of tradeoffs, and the simultaneous action of multiple mechanisms which are somewhat antagonistic to one another. This understanding can inform more effective team assembly and management strategies, enhancing innovation outcomes in applied settings.

References

Abernathy, William J., and Richard S. Rosenbloom. 1969. "Parallel Strategies in Development Projects." *Management Science* 15 (10): B-486.

Adair, Wendi L., Lindie H. Liang, and Ivona Hideg. 2017. "Buffering Against the Detrimental Effects of Demographic Faultlines: The Curious Case of Intragroup Conflict in Small Work Groups." *Negotiation and Conflict Management Research* 10 (1): 28–45. https://doi.org/10.1111/ncmr.12087.

Aggarwal, Ishani, and Anita Williams Woolley. 2019. "Team Creativity, Cognition, and Cognitive Style Diversity." *Management Science* 65 (4): 1586–99. https://doi.org/10.1287/mnsc.2017.3001.

Aggarwal, Ishani, Anita Williams Woolley, Christopher F. Chabris, and Thomas W. Malone. 2019. "The Impact of Cognitive Style Diversity on Implicit Learning in Teams." *Frontiers in Psychology* 10: 112.

Alexander, Christopher. 1964. *Notes on the Synthesis of Form*. Vol. 5. Harvard University Press.

Altonji, Joseph G., Erica Blom, and Costas Meghir. 2012. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers." *Annual Review of Economics* 4 (1): 185–223. https://doi.org/10.1146/annurev-economics-080511-110908.

Ancona, Deborah Gladstein, and David F. Caldwell. 1992. "Demography and Design: Predictors of New Product Team Performance." *Organization Science* 3 (3): 321–41.

Baldwin, Carliss Young, Kim B. Clark, and Kim B. Clark. 2000. *Design Rules: The Power of Modularity*. Vol. 1. MIT press.

Bell, Suzanne T., Anton J. Villado, Marc A. Lukasik, Larisa Belau, and Andrea L. Briggs. 2011. "Getting Specific about Demographic Diversity Variable and Team Performance Relationships: A Meta-Analysis." *Journal of Management* 37 (3): 709–43.

Bermiss, Sekou, Jeremiah Green, and John RM Hand. 2023. "Does Greater Diversity in Executive Race/Ethnicity Reliably Predict Better Future Firm Financial Performance?" *Journal of Economics, Race, and Policy*, 1–16.

Bezrukova, Katerina, Karen A. Jehn, Elaine L. Zanutto, and Sherry M. B. Thatcher. 2009. "Do Workgroup Faultlines Help or Hurt? A Moderated Model of Faultlines, Team Identification, and Group Performance." *Organization Science* 20 (1): 35–50. https://doi.org/10.1287/orsc.1080.0379.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb. 2020. "Are Ideas Getting Harder to Find?" *American Economic Review* 110 (4): 1104–44.

Boaler, Jo. 2016. *Mathematical Mindsets: Unleashing Students' Potential through Creative Math, Inspiring Messages and Innovative Teaching*. John Wiley & Sons.

Boone, Christophe, Woody Van Olffen, and Nadine Roijakkers. 2004. "Selection on the Road to a Career: Evidence of Personality Sorting in Educational Choice." *Journal of Career Development* 31 (1): 61–78. https://doi.org/10.1177/089484530403100105.

Bruns, Hille C. 2013. "Working Alone Together: Coordination in Collaboration across Domains of Expertise." *Academy of Management Journal* 56 (1): 62–83.

Brynjolfsson, Erik, Daniel Rock, and Prasanna Tambe. 2019. "How Will Machine Learning Transform the Labor Market?" *Governance in an Emerging New World* 619. https://www.hoover.org/research/how-will-machine-learning-transform-labor-market.

Bunderson, J. Stuart, and Kathleen M. Sutcliffe. 2002. "Comparing Alternative Conceptualizations of Functional Diversity in Management Teams: Process and Performance Effects." *Academy of Management Journal* 45 (5): 875–93.

Burke, Catriona M, and Michael J Morley. 2016. "On Temporary Organizations: A Review, Synthesis and Research Agenda." *Human Relations* 69 (6): 1235–58. https://doi.org/10.1177/0018726715610809.

Carton, Andrew M., and Jonathon N. Cummings. 2012. "A Theory of Subgroups in Work Teams." *Academy of Management Review* 37 (3): 441–70. https://doi.org/10.5465/amr.2009.0322.

———. 2013. "The Impact of Subgroup Type and Subgroup Configurational Properties on Work Team Performance." *Journal of Applied Psychology* 98 (5): 732.

Chiu, Yi-Te, and D. Sandy Staples. 2013. "Reducing Faultlines in Geographically Dispersed Teams: Self-Disclosure and Task Elaboration." *Small Group Research* 44 (5): 498–531. https://doi.org/10.1177/1046496413489735.

Cockburn, Iain M., Rebecca Henderson, and Scott Stern. 2018. "The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis." In *The Economics of Artificial Intelligence: An Agenda*, 115–46. University of Chicago Press. https://www.nber.org/system/files/chapters/c14006/c14006.pdf.

Corritore, Matthew, Amir Goldberg, and Sameer B. Srivastava. 2020. "Duality in Diversity: How Intrapersonal and Interpersonal Cultural Heterogeneity Relate to Firm Performance." *Administrative Science Quarterly* 65 (2): 359–94.

Cromwell, Johnathan R., Teresa M. Amabile, and Jean-François Harvey. 2018. "An Integrated Model of Dynamic Problem Solving within Organizational Constraints." In *Individual Creativity in the Workplace*, 53–81. Elsevier. https://www.sciencedirect.com/science/article/pii/B9780128132388000036.

Cummings, Anne, Jing Zhou, and Greg R. Oldham. 1993. "Demographic Differences and Employee Work Outcomes: Effects on Multiple Comparison Groups." In *Annual Meeting of the Academy of Management, Atlanta, GA*.

Cummings, Jonathon N., and Sara Kiesler. 2007. "Coordination Costs and Project Outcomes in Multi-University Collaborations." *Research Policy* 36 (10): 1620–34.

Dahlin, Kristina B., Laurie R. Weingart, and Pamela J. Hinds. 2005. "Team Diversity and Information Use." *Academy of Management Journal* 48 (6): 1107–23.

Davis, Jason P. 2016. "The Group Dynamics of Interorganizational Relationships: Collaborating with Multiple Partners in Innovation Ecosystems." *Administrative Science Quarterly* 61 (4): 621–61. https://doi.org/10.1177/0001839216649350.

De Jong, Bart, Nicole Gillespie, Ian Williamson, and Carol Gill. 2021. "Trust Consensus Within Culturally Diverse Teams: A Multistudy Investigation." *Journal of Management* 47 (8): 2135–68. https://doi.org/10.1177/0149206320943658.

Edmondson, Amy C., Richard M. Bohmer, and Gary P. Pisano. 2001. "Disrupted Routines: Team Learning and New Technology Implementation in Hospitals." *Administrative Science Quarterly* 46 (4): 685–716.

Elman, Cheryl, and M. O. Angela. 2007. "The Effects of Social Origins, Life Events, and Institutional Sorting on Adults' School Transitions." *Social Science Research* 36 (3): 1276–99.

Faraj, Samer, and Lee Sproull. 2000. "Coordinating Expertise in Software Development Teams." *Management Science* 46 (12): 1554–68.

Flache, Andreas, and Michael Mäs. 2008. "Why Do Faultlines Matter? A Computational Model of How Strong Demographic Faultlines Undermine Team Cohesion." *Simulation Modelling Practice and Theory* 16 (2): 175–91.

Fleming, L. 2001. "Recombinant Uncertainty in Technological Search." *Management Science* 47 (1): 117–32.

Fleming, Lee, and Olav Sorenson. 2001. "Technology as a Complex Adaptive System: Evidence from Patent Data." *Research Policy* 30 (7): 1019–39.

Fouarge, Didier, Ben Kriechel, and Thomas Dohmen. 2014. "Occupational Sorting of School Graduates: The Role of Economic Preferences." *Journal of Economic Behavior & Organization* 106: 335–51.

Garcia-Prieto, Patricia, Erwan Bellard, and Susan C. Schneider. 2003. "Experiencing Diversity, Conflict, and Emotions in Teams." *Applied Psychology* 52 (3): 413–40. https://doi.org/10.1111/1464-0597.00142.

Gibson, Cristina B., and Jennifer L. Gibbs. 2006. "Unpacking the Concept of Virtuality: The Effects of Geographic Dispersion, Electronic Dependence, Dynamic Structure, and National Diversity on Team Innovation." *Administrative Science Quarterly* 51 (3): 451–95.

Gibson, Cristina, and Freek Vermeulen. 2003. "A Healthy Divide: Subgroups as a Stimulus for Team Learning Behavior." *Administrative Science Quarterly* 48 (2): 202–39. https://doi.org/10.2307/3556657.

Goh, Kenneth T., Paul S. Goodman, and Laurie R. Weingart. 2013. "Team Innovation Processes: An Examination of Activity Cycles in Creative Project Teams." *Small Group Research* 44 (2): 159–94. https://doi.org/10.1177/1046496413483326.

Gruber, Marc, Dietmar Harhoff, and Karin Hoisl. 2013. "Knowledge Recombination Across Technological Boundaries: Scientists vs. Engineers." *Management Science* 59 (4): 837–51. https://doi.org/10.1287/mnsc.1120.1572.

Hamilton, Barton H, and Jack A Nickerson. 2003. "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation." *Journal of Political Economy*, 33.

Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan. 2012. "Diversity and Productivity in Production Teams." In *Advances in the Economic Analysis of Participatory & Labor-Managed Firms*, edited by Alex

Bryson, 13:99–138. Emerald Group Publishing Limited. https://doi.org/10.1108/S0885-3339(2012)0000013009.
Hansen, Zeynep K., Hideo Owan, and Jie Pan. 2006. "The Impact of Group Diversity on Performance and Knowledge Spillover–An Experiment in a College Classroom." National Bureau of Economic Research Cambridge, Mass., USA. https://www.nber.org/papers/w12251.
Harrison, David A., and Katherine J. Klein. 2007. "What's the Difference? Diversity Constructs as Separation, Variety, or Disparity in Organizations." *Academy of Management Review* 32 (4): 1199–1228.
Hill, Patrick L., Joshua J. Jackson, Nicole Nagy, Gabriel Nagy, Brent W. Roberts, Oliver Lüdtke, and Ulrich Trautwein. 2016. "Majoring in Selection, and Minoring in Socialization: The Role of the College Experience in Goal Change Post–High School." *Journal of Personality* 84 (2): 194–203. https://doi.org/10.1111/jopy.12151.
Hoever, Inga J., Daan Van Knippenberg, Wendy P. Van Ginkel, and Harry G. Barkema. 2012. "Fostering Team Creativity: Perspective Taking as Key to Unlocking Diversity's Potential." *Journal of Applied Psychology* 97 (5): 982.
Hong, Lu, and 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): 16385–89. https://doi.org/10.1073/pnas.0403723101.
Hoogendoorn, Sander, Simon C. Parker, and Mirjam Van Praag. 2017. "Smart or Diverse Start-up Teams? Evidence from a Field Experiment." *Organization Science* 28 (6): 1010–28.
Horwitz, Sujin K., and Irwin B. Horwitz. 2007. "The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography." *Journal of Management* 33 (6): 987–1015.
Hou, Tianyu, Julie Juan Li, and Jun Lin. 2021. "Recombination of Knowledge Components and Knowledge Impact: Neighboring Components versus Distant Components." *IEEE Transactions on Engineering Management*. https://ieeexplore.ieee.org/abstract/document/9597383/?casa_token=N8ZYS7X_RsQAAAAA:pYmSN71QHLijv_L5vAuK5tuKx_GPv4gS_utrzovDrjNAxLtqg3ewU0SnX1030Anth4xMs5CpGlg.
Huckman, Robert S., and Bradley R. Staats. 2011. "Fluid Tasks and Fluid Teams: The Impact of Diversity in Experience and Team Familiarity on Team Performance." *Manufacturing & Service Operations Management* 13 (3): 310–28.
Huckman, Robert S., Bradley R. Staats, and David M. Upton. 2009. "Team Familiarity, Role Experience, and Performance: Evidence from Indian Software Services." *Management Science* 55 (1): 85–100.
Humphrey, Stephen E., Federico Aime, Lily Cushenbery, Aaron D. Hill, and Joshua Fairchild. 2017. "Team Conflict Dynamics: Implications of a Dyadic View of Conflict for Team Performance." *Organizational Behavior and Human Decision Processes* 142: 58–70.
Jackson, Susan E., Aparna Joshi, and Niclas L. Erhardt. 2003. "Recent Research on Team and Organizational Diversity: SWOT Analysis and Implications." *Journal of Management* 29 (6): 801–30.
Jehn, Karen A., Katerina Bezrukova, and Sherry Thatcher. 2008. "Conflict, Diversity, and Faultlines in Workgroups." *The Psychology of Conflict and Conflict Management in Organizations* 6.
Jones, B. F. 2009. "The Burden of Knowledge and the 'Death of the Renaissance Man': Is Innovation Getting Harder?" *Review of Economic Studies* 76 (1): 283–317.
Jones, Benjamin F. 2009. "The Burden of Knowledge and the 'Death of the Renaissance Man': Is Innovation Getting Harder?" *The Review of Economic Studies* 76 (1): 283–317.
Joshi, Aparna. 2014. "By Whom and When Is Women's Expertise Recognized? The Interactive Effects of Gender and Education in Science and Engineering Teams." *Administrative Science Quarterly* 59 (2): 202–39. https://doi.org/10.1177/0001839214528331.
Joshi, Aparna, and Hyuntak Roh. 2009. "The Role of Context in Work Team Diversity Research: A Meta-Analytic Review." *Academy of Management Journal* 52 (3): 599–627.
Kaplan, Sarah, and Keyvan Vakili. 2015. "The Double-Edged Sword of Recombination in Breakthrough Innovation." *Strategic Management Journal* 36 (10): 1435–57.
Katila, Riitta. 2002. "New Product Search over Time: Past Ideas in Their Prime?" *Academy of Management Journal* 45 (5): 995–1010.

Kaur, Manpreet, and Hong Ren. 2023. "Positive Faultlines: An Unconventional Perspective on Team Compositional Dynamics." *Team Performance Management: An International Journal*. https://www.emerald.com/insight/content/doi/10.1108/TPM-03-2022-0025/full/html.

Kavadias, Stylianos, and Svenja C. Sommer. 2009. "The Effects of Problem Structure and Team Diversity on Brainstorming Effectiveness." *Management Science* 55 (12): 1899–1913.

Kelley, Tom. 2001. *The Art of Innovation: Lessons in Creativity from IDEO, America's Leading Design Firm*. Vol. 10. Currency. https://books.google.com/books?hl=en&lr=&id=ncaJDQAAQBAJ&oi=fnd&pg=PA1&dq=ideo+innovation+projects&ots=58Zh7BZfhG&sig=DblibKna-GMhNBjGKAEwM2-2ATE.

Korteling, JE (Hans), Gillian C. van de Boer-Visschedijk, Romy AM Blankendaal, Rudy C. Boonekamp, and Aletta R. Eikelboom. 2021. "Human-versus Artificial Intelligence." *Frontiers in Artificial Intelligence* 4: 622364.

Lakhani, Karim R., and Lars Bo Jeppesen. 2007. "Getting Unusual Suspects to Solve R&D Puzzles." *Harvard Business Review* 85: xxx–xxx.

Lau, Dora C., and J. Keith Murnighan. 1998. "Demographic Diversity and Faultlines: The Compositional Dynamics of Organizational Groups." *Academy of Management Review* 23 (2): 325–40.

———. 2005. "Interactions within Groups and Subgroups: The Effects of Demographic Faultlines." *Academy of Management Journal* 48 (4): 645–59.

Lazear, Edward P. 1999. "Globalisation and the Market for Team-Mates." *The Economic Journal* 109 (454): 15–40.

Leahey, Erin, Christine M. Beckman, and Taryn L. Stanko. 2017. "Prominent but Less Productive: The Impact of Interdisciplinarity on Scientists' Research." *Administrative Science Quarterly* 62 (1): 105–39.

Lifshitz-Assaf, Hila. 2018. "Dismantling Knowledge Boundaries at NASA: The Critical Role of Professional Identity in Open Innovation." *Administrative Science Quarterly* 63 (4): 746–82.

Lingo, Elizabeth Long, and Siobhán O'Mahony. 2010. "Nexus Work: Brokerage on Creative Projects." *Administrative Science Quarterly* 55 (1): 47–81. https://doi.org/10.2189/asqu.2010.55.1.47.

Lix, Katharina, Amir Goldberg, Sameer B. Srivastava, and Melissa A. Valentine. 2022. "Aligning Differences: Discursive Diversity and Team Performance." *Management Science* 68 (11): 8430–48. https://doi.org/10.1287/mnsc.2021.4274.

Lyons, Elizabeth. 2017. "Team Production in International Labor Markets: Experimental Evidence from the Field." *American Economic Journal: Applied Economics* 9 (3): 70–104. https://doi.org/10.1257/app.20160179.

Manata, Brian, Angelo Joseph Garcia, Sinem Mollaoglu, and Vernon D. Miller. 2021. "The Effect of Commitment Differentiation on Integrated Project Delivery Team Dynamics: The Critical Roles of Goal Alignment, Communication Behaviors, and Decision Quality." *International Journal of Project Management* 39 (3): 259–69.

Marion, Tucker James, and Marc H. Meyer. 2018. "Organizing to Achieve Modular Architecture across Different Products." *IEEE Transactions on Engineering Management* 65 (3): 404–16.

Marx, Benjamin, Vincent Pons, and Tavneet Suri. 2021. "Diversity and Team Performance in a Kenyan Organization☆." *Journal of Public Economics* 197: 104332.

Mäs, Michael, Andreas Flache, Károly Takács, and Karen A. Jehn. 2013. "In the Short Term We Divide, in the Long Term We Unite: Demographic Crisscrossing and the Effects of Faultlines on Subgroup Polarization." *Organization Science* 24 (3): 716–36. https://doi.org/10.1287/orsc.1120.0767.

Mathieu, J. E., S. I. Tannenbaum, J. S. Donsbach, and G. M. Alliger. 2013. "Achieving Optimal Team Composition for Success." *Developing and Enhancing High-Performance Teams: Evidence-Based Practices and Advice*, 520–51.

Mayo, Anna T. 2022. "Syncing Up: A Process Model of Emergent Interdependence in Dynamic Teams." *Administrative Science Quarterly* 67 (3): 821–64. https://doi.org/10.1177/00018392221096451.

Mayo, Anna T., Anita Williams Woolley, and Rosalind M. Chow. 2020. "Unpacking Participation and Influence: Diversity's Countervailing Effects on Expertise Use in Groups." *Academy of Management Discoveries* 6 (2): 300–319.

Meyer, Bertolt, Andreas Glenz, Mirko Antino, Ramón Rico, and Vicente González-Romá. 2014. "Faultlines and Subgroups: A Meta-Review and Measurement Guide." *Small Group Research* 45 (6): 633–70. https://doi.org/10.1177/1046496414552195.

Meyer, Bertolt, Meir Shemla, Jia Li, and Jürgen Wegge. 2015. "On the Same Side of the Faultline: Inclusion in the Leader's Subgroup and Employee Performance." *Journal of Management Studies* 52 (3): 354–80. https://doi.org/10.1111/joms.12118.

Murnighan, John K., and Daniel J. Brass. 1991. "Intraorganizational Coalitions." In *The Handbook of Negotiation Research*, 283–306. JAI Press. https://www.scholars.northwestern.edu/en/publications/intraorganizational-coalitions.

Murnighan, Keith, and Dora Lau. 2017. "Faultlines." In *Oxford Research Encyclopedia of Business and Management*. https://oxfordre.com/business/display/10.1093/acrefore/9780190224851.001.0001/acrefore-9780190224851-e-60?rskey=Cxl2mv.

Nagle, Frank, and Florenta Teodoridis. 2020. "Jack of All Trades and Master of Knowledge: The Role of Diversification in New Distant Knowledge Integration." *Strategic Management Journal* 41 (1): 55–85. https://doi.org/10.1002/smj.3091.

Page, Scott E. 2019. *The Diversity Bonus: How Great Teams Pay off in the Knowledge Economy*. Princeton University Press. https://books.google.com/books?hl=en&lr=&id=oz9xDwAAQBAJ&oi=fnd&pg=PP11&dq=page+diversity&ots=uCRLpruMrJ&sig=mxwmqfiOWpMxdGPV_f8SmlT7eFU.

Parnas, David Lorge. 1972. "On the Criteria to Be Used in Decomposing Systems into Modules." *Communications of the ACM* 15 (12): 1053–58.

Pascarella, Ernest T., and Patrick T. Terenzini. 2005. *How College Affects Students: A Third Decade of Research. Volume 2*. ERIC.

Paulus, Paul B., Karen I. Van Der Zee, and Jared Kenworthy. 2016. "Cultural Diversity and Team Creativity." In *The Palgrave Handbook of Creativity and Culture Research*, edited by Vlad Petre Glăveanu, 57–76. London: Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-46344-9_4.

Pelled, Lisa Hope. 1996. "Demographic Diversity, Conflict, and Work Group Outcomes: An Intervening Process Theory." *Organization Science* 7 (6): 615–31.

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

Perretti, Fabrizio, and Giacomo Negro. 2007. "Mixing Genres and Matching People: A Study in Innovation and Team Composition in Hollywood." *Journal of Organizational Behavior* 28 (5): 563–86. https://doi.org/10.1002/job.464.

Pollok, Patrick, André Amft, Kathleen Diener, Dirk Lüttgens, and Frank T. Piller. 2021. "Knowledge Diversity and Team Creativity: How Hobbyists Beat Professional Designers in Creating Novel Board Games." *Research Policy*, January, 104174. https://doi.org/10.1016/j.respol.2020.104174.

Polzer, J. T., E. A. Mannix, and M. A. Neale. 1998. "INTEREST ALIGNMENT AND COALITIONS IN MULTIPARTY NEGOTIATION." *Academy of Management Journal* 41 (1): 42–54. https://doi.org/10.2307/256896.

Polzer, Jeffrey T., Laurie P. Milton, and William B. Swarm Jr. 2002. "Capitalizing on Diversity: Interpersonal Congruence in Small Work Groups." *Administrative Science Quarterly* 47 (2): 296–324.

Privman, Robin, Starr Roxanne Hiltz, and Yiran Wang. 2013. "In-Group (Us) versus out-Group (Them) Dynamics and Effectiveness in Partially Distributed Teams." *IEEE Transactions on Professional Communication* 56 (1): 33–49.

Reagans, Ray, Ezra Zuckerman, and Bill McEvily. 2004. "How to Make the Team: Social Networks vs. Demography as Criteria for Designing Effective Teams." *Administrative Science Quarterly* 49 (1): 101–33.

Ren, Yuqing, Jilin Chen, and John Riedl. 2016. "The Impact and Evolution of Group Diversity in Online Open Collaboration." *Management Science* 62 (6): 1668–86.

Rico, Ramón, Miriam Sánchez-Manzanares, Mirko Antino, and Dora Lau. 2012. "Bridging Team Faultlines by Combining Task Role Assignment and Goal Structure Strategies." *Journal of Applied Psychology* 97 (2): 407.

Roberson, Quinetta M. 2019a. "Diversity in the Workplace: A Review, Synthesis, and Future Research Agenda." *Annual Review of Organizational Psychology and Organizational Behavior* 6: 69–88.

———. 2019b. “Diversity in the Workplace: A Review, Synthesis, and Future Research Agenda.” *Annual Review of Organizational Psychology and Organizational Behavior* 6 (1): 69–88. https://doi.org/10.1146/annurev-orgpsych-012218-015243.

Robinson, Peter M. 1988. “Root-N-Consistent Semiparametric Regression.” *Econometrica: Journal of the Econometric Society*, 931–54.

Roksa, Josipa, and Tania Levey. 2010. “What Can You Do with That Degree? College Major and Occupational Status of College Graduates over Time.” *Social Forces* 89 (2): 389–415.

Rosendahl Huber, Laura, Randolph Sloof, Mirjam Van Praag, and Simon C. Parker. 2020. “Diverse Cognitive Skills and Team Performance: A Field Experiment Based on an Entrepreneurship Education Program.” *Journal of Economic Behavior & Organization* 177 (September): 569–88. https://doi.org/10.1016/j.jebo.2020.06.030.

Schölmerich, Franziska, Carsten C. Schermuly, and Jürgen Deller. 2016. “How Leaders’ Diversity Beliefs Alter the Impact of Faultlines on Team Functioning.” *Small Group Research* 47 (2): 177–206. https://doi.org/10.1177/1046496416630960.

Schumpeter, Joseph A. 1942. “Capitalism, Socialism and Democracy.” *New York*.

Simon, H A. 1991. “Bounded Rationality and Organizational Learning.” *Organization Science* 2 (1): 125–34.

Simon, Herbert A. 1991. “The Architecture of Complexity.” In *Facets of Systems Science*, 457–76. Springer.

Simpson, Timothy W., Tucker Marion, Olivier De Weck, Katja Ho\" ltta\"-Otto, Michael Kokkolaras, and Steve B. Shooter. 2006. “Platform-Based Design and Development: Current Trends and Needs in Industry.” In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 4255:801–10.

Smith-Doerr, Laurel, Sharla N. Alegria, and Timothy Sacco. 2017. “How Diversity Matters in the US Science and Engineering Workforce: A Critical Review Considering Integration in Teams, Fields, and Organizational Contexts.” *Engaging Science, Technology, and Society* 3: 139–53.

Sternitzke, Christian. 2009. “Patents and Publications as Sources of Novel and Inventive Knowledge.” *Scientometrics* 79 (3): 551–61.

Taylor, a., and H R Greve. 2006. “Superman or the Fantastic Four? Knowledge Combination and Experience in Innovative Teams.” *Academy of Management Journal* 49 (4): 723–40.

Teodoridis, Florenta. 2018. “Understanding Team Knowledge Production: The Interrelated Roles of Technology and Expertise.” *Management Science* 64 (8): 3625–48. https://doi.org/10.1287/mnsc.2017.2789.

Teodoridis, Florenta, Michaël Bikard, and Keyvan Vakili. 2019. “Creativity at the Knowledge Frontier: The Impact of Specialization in Fast- and Slow-Paced Domains.” *Administrative Science Quarterly* 64 (4): 894–927. https://doi.org/10.1177/0001839218793384.

Thatcher, Sherry MB, Karen A. Jehn, and Elaine Zanutto. 2003. “Cracks in Diversity Research: The Effects of Diversity Faultlines on Conflict and Performance.” *Group Decision and Negotiation* 12: 217–41.

Thatcher, Sherry, and Pankaj C. Patel. 2011. “Demographic Faultlines: A Meta-Analysis of the Literature.” *Journal of Applied Psychology* 96 (6): 1119.

Uzzi, B., S. Mukherjee, M. Stringer, and B. Jones. 2013. “Atypical Combinations and Scientific Impact.” *Science* 342 (6157): 468–72. https://doi.org/10.1126/science.1240474.

Uzzi, Brian, and Jarrett Spiro. 2005. “Collaboration and Creativity: The Small World Problem.” *American Journal of Sociology* 111 (2): 447–504. https://doi.org/10.1086/432782.

Vakili, Keyvan, and Sarah Kaplan. 2019. “Organizing for Innovation: How Team Configurations Vary With Modularity and Breadth of Application.” *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3509461.

———. 2021. “Organizing for Innovation: A Contingency View on Innovative Team Configuration.” *Strategic Management Journal* 42 (6): 1159–83. https://doi.org/10.1002/smj.3264.

Van Der Vegt, Gerben S., and J. Stuart Bunderson. 2005. “Learning and Performance in Multidisciplinary Teams: The Importance of Collective Team Identification.” *Academy of Management Journal* 48 (3): 532–47. https://doi.org/10.5465/amj.2005.17407918.

Van Dijk, Hans, Marloes L. Van Engen, and Daan Van Knippenberg. 2012. “Defying Conventional Wisdom: A Meta-Analytical Examination of the Differences between Demographic and Job-Related Diversity Relationships with Performance.” *Organizational Behavior and Human Decision Processes* 119 (1): 38–53.

Van Knippenberg, Daan, S. Alexander Haslam, and Michael J. Platow. 2007. "Unity through Diversity: Value-in-Diversity Beliefs, Work Group Diversity, and Group Identification." *Group Dynamics: Theory, Research, and Practice* 11 (3): 207.

Wang, Dawei, Liping Zhu, Phil Maguire, Yixin Liu, Kaiyuan Pang, Zhenying Li, and Yixin Hu. 2016. "The Influence of Social Comparison and Peer Group Size on Risky Decision-Making." *Frontiers in Psychology* 7: 1232.

Weisberg, Robert W. 1999. "Creativity and Knowledge: A Challenge to Theories." https://psycnet.apa.org/record/1998-08125-012.

Weitzman, Martin L. 1998. "Recombinant Growth." *The Quarterly Journal of Economics* 113 (2): 331–60.

Wiswall, Matthew, and Basit Zafar. 2021. "New Approaches to Understanding Choice of Major." *NBER Reporter*, no. 2: 18–21.

Woolley, Anita Williams, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone. 2010. "Evidence for a Collective Intelligence Factor in the Performance of Human Groups." *Science* 330 (6004): 686–88. https://doi.org/10.1126/science.1193147.

Woolley, Anita Williams, Margaret E. Gerbasi, Christopher F. Chabris, Stephen M. Kosslyn, and J. Richard Hackman. 2008. "Bringing in the Experts: How Team Composition and Collaborative Planning Jointly Shape Analytic Effectiveness." *Small Group Research* 39 (3): 352–71.

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi. 2007a. "The Increasing Dominance of Teams in Production of Knowledge." *Science* 316 (5827): 1036–39.

———. 2007b. "The Increasing Dominance of Teams in Production of Knowledge." *Science* 316 (5827): 1036–39.

Xiao, Ting, Mona Makhija, and Samina Karim. 2022. "A Knowledge Recombination Perspective of Innovation: Review and New Research Directions." *Journal of Management* 48 (6): 1724–77.

Yang, Yang, Tanya Y. Tian, Teresa K. Woodruff, Benjamin F. Jones, and Brian Uzzi. 2022. "Gender-Diverse Teams Produce More Novel and Higher-Impact Scientific Ideas." *Proceedings of the National Academy of Sciences* 119 (36): e2200841119.

Yayavaram, Sai, and Gautam Ahuja. 2008. "Decomposability in Knowledge Structures and Its Impact on the Usefulness of Inventions and Knowledge-Base Malleability." *Administrative Science Quarterly* 53 (2): 333–62.

Yilmaz, Gamze, and Jorge Peña. 2014. "The Influence of Social Categories and Interpersonal Behaviors on Future Intentions and Attitudes to Form Subgroups in Virtual Teams." *Communication Research* 41 (3): 333–52. https://doi.org/10.1177/0093650212443696.

Zenger, T. R., and B. S. Lawrence. 1989. "ORGANIZATIONAL DEMOGRAPHY: THE DIFFERENTIAL EFFECTS OF AGE AND TENURE DISTRIBUTIONS ON TECHNICAL COMMUNICATION." *Academy of Management Journal* 32 (2): 353–76. https://doi.org/10.2307/256366.

Zhang, Lei, and Isin Guler. 2020. "How to Join the Club: Patterns of Embeddedness and the Addition of New Members to Interorganizational Collaborations." *Administrative Science Quarterly* 65 (1): 112–50. https://doi.org/10.1177/0001839219834011.

FIGURES

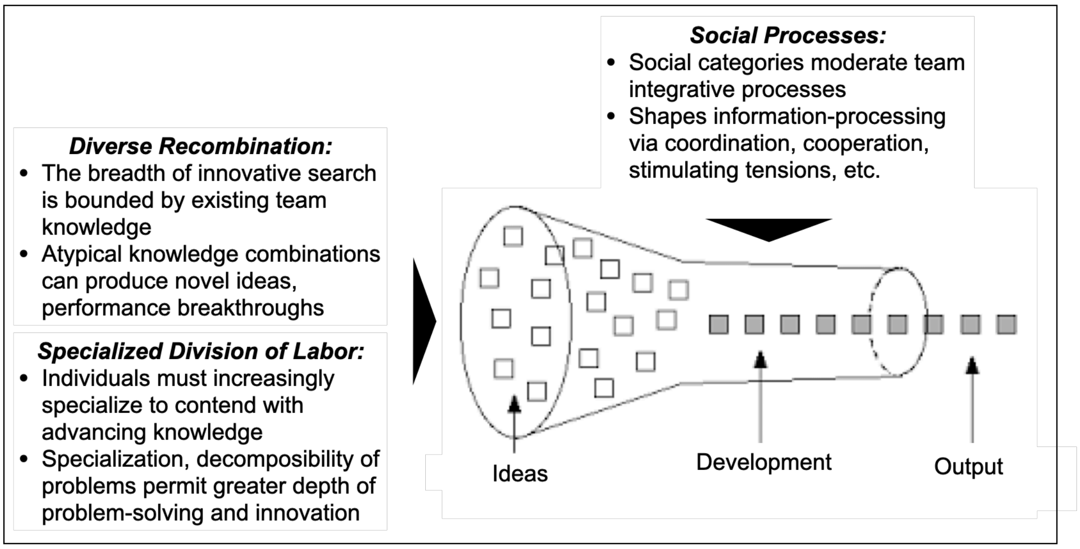

Figure 1 Knowledge-as-Inputs vs. Knowledge-as-Moderator of Innovation Processes

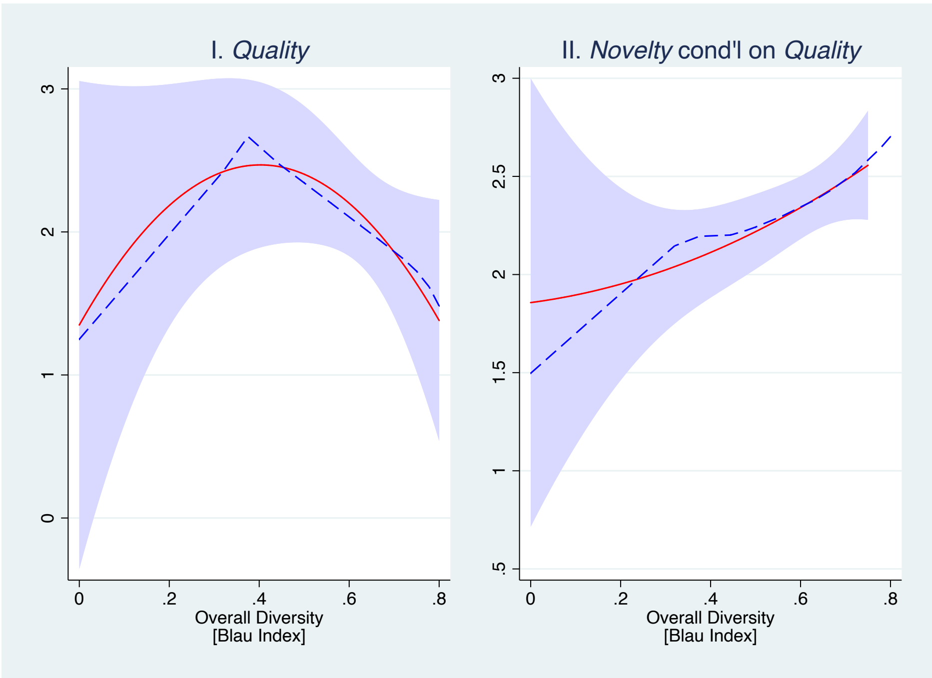

Figure 2 Experimental Variation in Knowledge Diversity & Innovation Outcomes

Red – Quadratic Estimate; Red – Non-Parametric Estimate

Notes. The red line is a quadratic model estimated with OLS. The non-parametric estimate conditional mean appears as the dashed blue line, with 90 percent confidence intervals, and uses a locally weighted linear model with an Epanechnikov weighting kernel estimated by a procedure described by Robinson (1988). No. obs. = 218 teams (872 individuals).

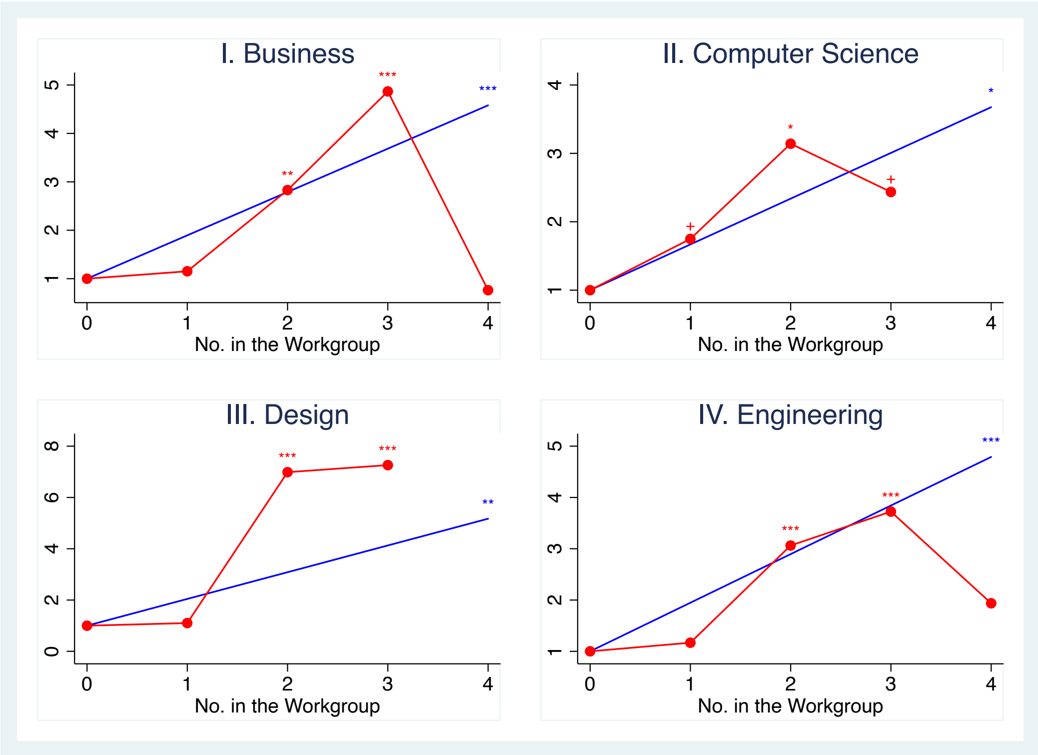

Figure 3 Non-Linear Effects of Adding Team Members on Overall Innovation *Quality*

Blue – Linear Estimates, Table 3 model (5)
Red – Estimates of Individual Dummies for Different Counts

Notes. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. The red dots indicate point estimates from a re-estimation of Table 3 model (5) in which the linear controls for counts of team members from related fields are replaced by a set of dummies for 1, 2, 3, or 4 for each of these 4 fields. Counts from unrelated fields continue to be specified as linear. No. obs. = 218 teams (872 individuals).

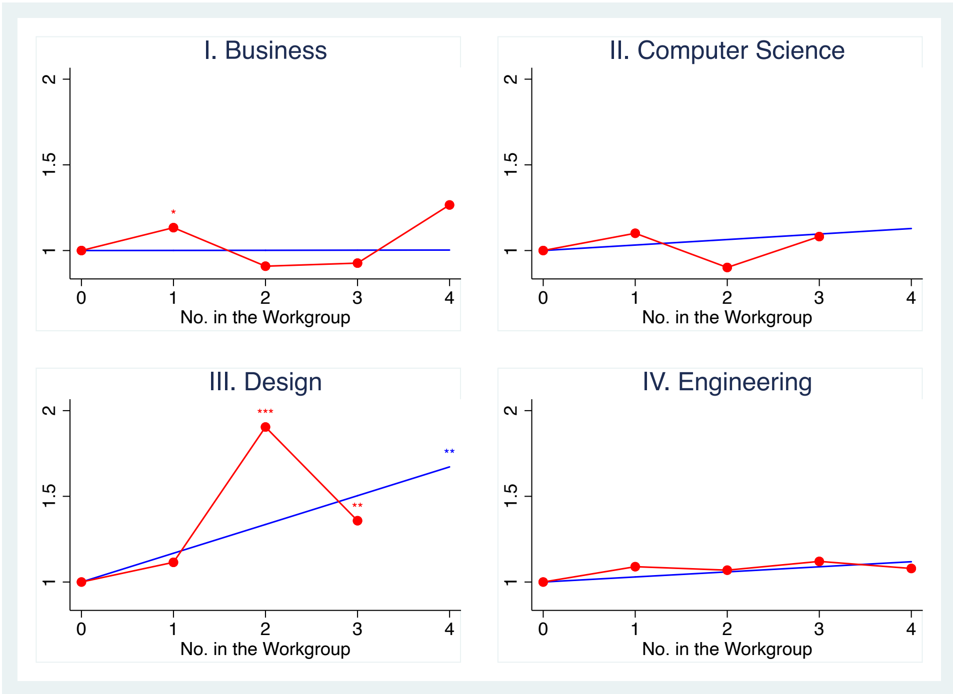

Figure 4 Non-Linear Effects of Adding Team Members on Innovation *Novelty* (Controlling for *Quality*)

Blue – Linear Estimates, Table 3 model (9)
Red – Estimates of Individual Dummies for Different Counts

Notes. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. The red dots indicate point estimates from a re-estimation of Table 3 model (9) in which the linear controls for counts of team members from related fields are replaced by a set of dummies for 1, 2, 3, or 4 for each of these 4 fields. Counts from unrelated fields continue to be specified as linear. No. obs. = 218 teams (872 individuals).

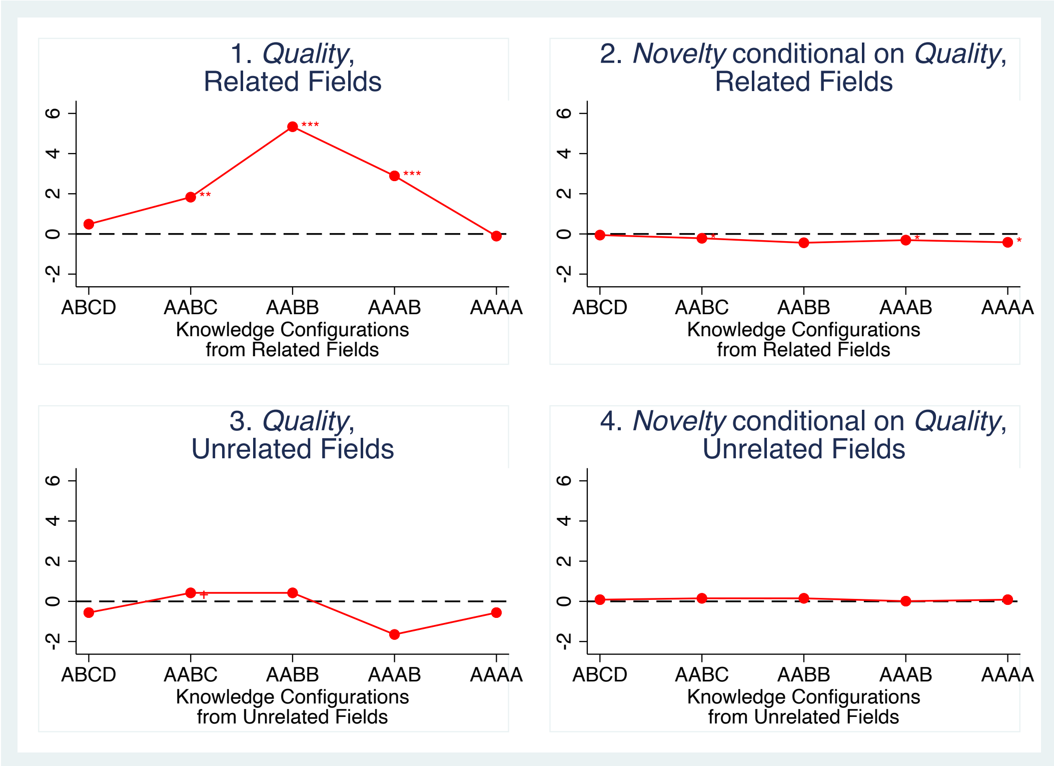

Figure 5 Effects of Subgroups on Overall Innovation *Quality*

Notes. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. Each of the panels represents estimates from different model estimates, with either Quality or Novelty as dependent variables. Estimates include a full set of controls. In addition, *Novelty* estimates control for *Quality* in order to test whether *Novelty* responds to to differences in subgroups and knowledge configurations than Quality. Separate models were used to estimate the coefficients on subgroups for related and unrelated fields. Estimating them within the same model does not statistically alter results, but results in less precise estimates. No. obs. = 218 teams (872 individuals).

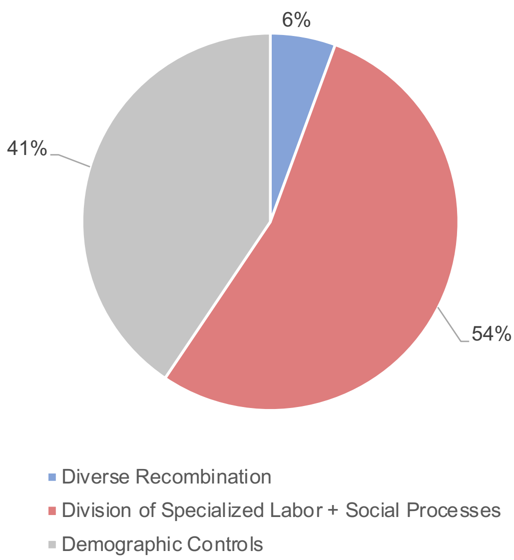

Figure 6 Breakdown of the Share of Variation Explained

TABLES

Table 1 Summary of 3 Theoretical Perspectives and their Alternate Predictions

Theoretical Perspective	Relevant Types of Knowledge Differences	Predictions for Innovation Performance	Implications for Team Design	E.g.
Diverse Recombination	Distinct and possibly atypical knowledge combinations	Varied and atypical knowledge combinations lead to expanded search for novel solutions, and possibly higher performance	The optimal team draws together some degree of atypical knowledge to provoke greater exploration	(Lee Fleming and Sorenson 2001; B. Uzzi et al. 2013; Kaplan and Vakili 2015)
Division of Specialized Labor	Knowledge expertise mapped to the nature and structure (decomposibility) of the problem at hand	Deep expertise matched to components of a problem enable superior solutions	The optimal team knowledge precisely spans the scope of the problem, with specialization mapping to problem decomposibility	(Jones 2009; Wuchty, Jones, and Uzzi 2007b; H A Simon 1991)
Social Processes	Differences in knowledge, training, and expertise that reflect differences in sociology groups and identities	Greater diversity and divisions produce greater frictions, subgroups and faultiness are more relevant than diversity per se in moderating team effectiveness	The optimal team composition will be that which enhance the social dynamics and interactions that enhance team integration of knowledge on whether they have positive or negative effect on performance	(Pelled, Eisenhardt, and Xin 1999; Lau and Murnighan 2005; Bunderson and Sutcliffe 2002)

Table 2 Descriptive Statistics

	Mean	Std. Dev.	Min.	Max.
Innovation Performance				
Quality	2.10	3.00	0	10
Novelty	2.01	2.89	0	9
Knowledge:				
Business	0.78	0.79	0	4
Computer Science	0.44	0.61	0	3
Design	0.25	0.49	0	3
Engineering	1.22	0.93	0	4
Health & Nursing	0.17	0.41	0	2
Humanities	0.55	0.70	0	3
Law	0.05	0.21	0	1
Science	0.33	0.58	0	3
Blau Index	0.57	0.16	0	0.8
Graduate	0.76	0.77	0	4
Demographics:				
Female	1.48	0.94	0	4
Avg Age	30.5	5.0	21.7	48
Race = "Black"	0.24	0.49	0	3
Race = "Asia-Pacific"	1.56	0.96	0	4
Race = "Hispanic"	0.18	0.40	0	2
Race = "White"	1.83	0.99	0	4

Notes. No. obs. = 218 teams of 4 (872 individuals). 74 teams (34%) successfully submitted completed solutions and 144 teams did not. For teams successfully submitting, mean Quality was 6.13 (std. dev. = 1.21) and mean Novelty was 5.87 (std. dev. = 1.27). (Section 5.3 includes analysis of probability of successfully submitting versus outcomes conditional on successful submission.)

Table 3 The Effect of Varying Counts of Knowledge Types on Innovation Performance

Dep. Var.:	*Quality*					*Novelty*			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Team Member Counts by Knowledge Fields									
Counts for Related Fields:									
Business	1.16***	1.16***	1.12***	1.03***	1.01***	1.13***	0.99***	0.03	0.03
	(0.34)	(0.34)	(0.35)	(0.35)	(0.35)	(0.32)	(0.33)	(0.05)	(0.05)
Computer Science	0.69*	0.70*	0.72*	0.78*	0.80**	0.73*	0.83**	0.07	0.07
	(0.40)	(0.40)	(0.41)	(0.41)	(0.41)	(0.39)	(0.39)	(0.05)	(0.06)
Design	1.30***	1.20**	1.29***	1.30***	1.20**	1.42***	1.35***	0.18***	0.21***
	(0.47)	(0.50)	(0.47)	(0.46)	(0.49)	(0.47)	(0.48)	(0.06)	(0.07)
Engineering	0.85***	0.88***	0.83***	0.96***	1.05***	0.87***	1.06***	0.06	0.06
	(0.31)	(0.31)	(0.31)	(0.32)	(0.32)	(0.30)	(0.32)	(0.04)	(0.05)
Counts for Unrelated Fields:									
Health & Nursing	0.83	0.79	0.76	0.63	0.61	0.91	0.69	0.11	0.11
	(0.53)	(0.54)	(0.53)	(0.51)	(0.54)	(0.56)	(0.54)	(0.11)	(0.10)
Humanities	0.49	0.41	0.45	0.18	0.07	0.51	0.11	0.04	0.05
	(0.36)	(0.36)	(0.36)	(0.39)	(0.40)	(0.34)	(0.39)	(0.05)	(0.06)
Law	-0.05	0.08	-0.01	-0.02	0.29	-0.13	0.17	-0.08	-0.10
	(0.78)	(0.75)	(0.78)	(0.87)	(0.83)	(0.71)	(0.78)	(0.10)	(0.10)
Sciences	0.43	0.38	0.38	0.28	0.22	0.41	0.23	0.01	0.02
	(0.35)	(0.36)	(0.36)	(0.36)	(0.36)	(0.34)	(0.35)	(0.05)	(0.06)
Grad Degree	0.09	0.05	0.04	0.20	0.20	0.12	0.25	0.04*	0.05**
	(0.13)	(0.14)	(0.15)	(0.14)	(0.17)	(0.12)	(0.16)	(0.02)	(0.02)
Controls									
Gender		Y			Y		Y		Y
Age			Y		Y		Y		Y
Race				Y	Y		Y		Y
Quality								0.95***	0.95***
								(0.01)	(0.01)
Constant	-1.32	-1.59	-2.98	-1.77	-0.94	-1.56	-.83	-0.31**	.06
	(0.98)	(0.99)	(1.88)	(1.51)	(2.68)	(0.96)	(2.51)	(0.15)	(0.29)
Adjusted-R^2	0.03	0.03	0.02	0.10	0.10	0.04	0.11	0.98	0.98

Notes. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. OLS model coefficient estimates; robust standard errors in parentheses; gender controls = No. of women; age controls = average and range of ages; race controls = counts (see Section 3.4 discussion). No. obs. = 218 teams (872 individuals).

Table 4 The Effect of Knowledge Configurations & Subgroups on Innovation Performance

Dep. Var.:	*Quality*			*Novelty*					
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Knowledge Configurations & Subgroups (Related Fields)									
2 from Same Field: AA		1.34**	1.40***		1.13*	1.24***		-0.16	-0.10
		(0.60)	(0.47)		(0.58)	(0.47)		(0.11)	(0.10)
2 from One Field & 2 from Another: AABB		3.51***	3.62***		3.13***	3.31***		-0.22	-0.16
		(0.91)	(0.92)		(0.92)	(0.97)		(0.22)	(0.26)
3 from Same Field AAAB		2.40**	2.44***		2.05**	2.19***		-0.25	-0.15
		(0.93)	(0.72)		(0.88)	(0.70)		(0.16)	(0.12)
4 from Same Field: AAAA		-0.59	-0.62		-.93	-0.82		-0.36	-0.22
		(1.05)	(0.68)		(1.03)	(0.68)		(0.23)	(0.14)
Team Member Counts									
Counts for Related Fields:									
Business	1.01***	0.31		0.99***	0.40		0.03	0.11*	
	(0.35)	(0.39)		(0.33)	(0.38)		(0.05)	(0.06)	
Computer Science	0.80**	0.56		0.83**	0.63		0.07	0.09	
	(0.41)	(0.42)		(0.39)	(0.40)		(0.06)	(0.06)	
Design	1.20**	0.70		1.35***	0.92**		0.21***	0.25***	
	(0.49)	(0.46)		(0.48)	(0.45)		(0.07)	(0.07)	
Engineering	1.05***	0.30		1.06***	0.43		0.06	0.15**	
	(0.32)	(0.44)		(0.32)	(0.42)		(0.05)	(0.07)	
Counts for Unrelated Fields	Y	Y	Y	Y	Y	Y	Y	Y	Y
Grad Degree	0.20	0.23	0.15	0.25	0.28*	0.18	0.05**	0.05**	0.03
	(0.17)	(0.17)	(0.16)	(0.16)	(0.16)	(0.15)	(0.02)	(0.02)	(0.02)
Controls									
Gender	Y	Y	Y	Y	Y	Y	Y	Y	Y
Age	Y	Y	Y	Y	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y	Y	Y	Y	Y
Quality							0.95***	0.95***	0.95***
							(0.02)	(0.02)	(0.02)
Constant	-0.94	0.49	-0.19	-0.83	0.41	-0.29	0.06	-0.06	-0.11
	(2.68)	(2.74)	(2.73)	(2.51)	(2.57)	(2.58)	(0.29)	(0.31)	(0.35)
Adjusted-R^2	0.10	0.17	0.18	0.11	0.17	0.17	0.98	0.98	0.98

Notes. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. OLS model coefficient estimates; robust standard errors in parentheses; gender controls = No. of women; age controls = average and range of ages; race controls = counts (see Section 3.4 discussion). No. obs. = 218 teams (872 individuals).

Table 5 The Effect of Gender-Balanced Teams on Innovation Performance

Dep. Var.:	*Quality*				*Novelty*	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Team Gender Composition						
WMMM	0.42	0.43	0.89*	0.90*	0.87*	0.01
	(0.57)	(0.60)	(0.50)	(0.50)	(0.49)	(0.07)
(Balanced) WWMM	1.19*	1.26*	1.92***	2.09***	2.04***	0.04
	(0.60)	(0.65)	(0.53)	(0.53)	(0.52)	(0.08)
WWWM	0.60	0.65	1.25*	1.42**	1.29**	-0.06
	(0.73)	(0.83)	(0.69)	(0.64)	(0.60)	(0.10)
WWWW	-1.49***	-0.87				
	(0.47)	(0.86)				
Knowledge Configurations & Subgroups (Related Fields)						
2 from Same Field				1.31**	1.09*	-0.16
				(0.60)	(0.57)	(0.11)
(Balanced) 2 from Same Field, 2 From Another				3.54***	3.16***	-0.22
				(0.90)	(0.90)	(0.22)
3 from Same Field				2.41**	2.05**	-0.25
				(0.95)	(0.91)	(0.16)
4 from Same Field				-1.00	-1.34	-0.39*
				(1.08)	(1.03)	(0.22)
Team Member Counts						
Counts by Related Knowledge		Y	Y	Y	Y	Y
Counts by Unrelated Knowledge Fields		Y	Y	Y	Y	Y
Grad Degree		Y	Y	Y	Y	Y
Controls						
Age			Y	Y	Y	Y
Race			Y	Y	Y	Y
Quality						0.95***
						0.01
Constant	0.45	0.40	0.19	0.40	-0.06	-0.09
	(2.82)	(2.79)	(2.94)	(2.80)	(2.69)	(0.31)
Adjusted-R^2	0.01	0.03	0.12	0.19	0.19	0.98

Notes. W and M indicate the mix of women and men; *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. OLS model coefficient estimates; robust standard errors in parentheses; age controls = average and range of ages; race controls = counts (see Section 3.4 discussion). No. obs. = 218 teams (872 individuals).

Table 6 The Effect of Knowledge Subgroups on the Probability of Completing a Submission

Dep. Var.:	*Successful Completion*	*Quality | Complete*	*Novelty | Complete*	*Novelty | Quality, Complete*
Model:	(1)	(2)	(3)	(4)
Knowledge Configurations & Subgroups (Related Fields)				
2 from Same Field: AA	0.18*	0.64*	0.20	-0.30
	(0.10)	(0.34)	(0.32)	(0.26)
2 from One Field & 2 from	0.50***	0.27	-0.28	-0.49
	(0.14)	(0.49)	(0.64)	(0.34)
3 from Same Field AAAB	0.34**	0.44	-0.53	-0.87**
	(0.14)	(0.58)	(0.52)	(0.36)
4 from Same Field: AAAA	-0.16			
	(0.16)			
Team Member Counts				
Counts for Related Fields				
Business	0.05	0.21	0.57	0.40
	(0.06)	(0.40)	(0.47)	(0.31)
Computer Science	0.08	0.02	0.31	0.30
	(0.07)	(0.33)	(0.39)	(0.25)
Design	0.10	0.29	0.92**	0.69***
	(0.07)	(0.29)	(0.37)	(0.25)
Engineering	0.07	-0.24	0.24	0.43
	(0.07)	(0.39)	(0.45)	(0.31)
Counts for Unrelated Fields	Y	Y	Y	Y
Grad Degree	0.03	0.14	0.30**	0.19**
	(0.03)	(0.13)	(0.13)	(0.08)
Controls				
Gender	Y	Y	Y	Y
Age	Y	Y	Y	Y
Race	Y	Y	Y	Y
Quality				0.78***
				(0.09)
Constant	0.04	5.91*	6.13**	1.50
	(0.43)	(3.23)	(2.86)	(1.77)
R^2	0.12	0.11	0.20	0.69

Notes. 35 percent of teams successfully completed a submission. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. OLS model coefficient estimates; robust standard errors in parentheses; gender controls = No. of women; age controls = average and range of ages; race controls = counts (see Section 3.4 discussion). No. obs. = 218 teams (872 individuals) in model (1); No. obs. = 74 teams in models (2), (3), and (4), as 74 teams, or 35 percent, successfully made completed submissions.

Table 7 Effects of Diversity per se on Innovation Outcomes

Dep. Var.:	*Quality*		*Novelty*	
Model:	(1)	(2)	(3)	(4)
Diversity per-se				
Blau Index	3.68	3.49**	3.87**	0.53**
	(2.75)	(1.76)	(1.63)	(0.26)
Knowledge Configurations & Subgroups (Related Fields)				
2 from Same Field: AA	2.02**	1.72***	1.62***	-0.03
	(0.80)	(0.46)	(0.45)	(0.08)
2 from One Field & 2	4.02***	4.01***	3.78***	-0.06
from Another: AABB	(1.01)	(0.88)	(0.94)	(0.26)
3 from Same Field AAAB	4.12***	3.61***	3.53***	0.07
	(1.54)	(0.77)	(0.75)	(0.12)
4 from Same Field: AAAA	2.48	1.65	1.76	0.18
	(2.57)	(1.19)	(1.12)	(0.19)
Team Member Counts				
Counts for Related Fields				
Business	-0.20			
	(0.55)			
Computer Science	0.06			
	(0.55)			
Design	0.17			
	(0.59)			
Engineering	-0.30			
	(0.60)			
Counts for Unrelated Fields				
Health & Nursing	0.17			
	(0.56)			
Humanities	-0.15			
	(0.45)			
Law	0.12			
	(0.77)			
Sciences	-0.09			
	(0.38)			
Grad Degree	0.26	0.24	0.28*	0.05*
	(0.16)	(0.17)	(0.16)	(0.02)
Controls				
Gender	Y	Y	Y	Y
Age	Y	Y	Y	Y
Race	Y	Y	Y	Y
Quality				0.96***
				(0.01)
Constant	0.19	-0.03	-0.09	-0.06
	(2.68)	(2.61)	(2.46)	(0.35)
Adjusted-R^2	0.18	0.20	0.19	0.98

Notes. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. OLS model coefficient estimates; robust standard errors in parentheses; gender controls = No. of women; age controls = average and range of ages; race controls = counts (see Section 3.4 discussion). No. obs. = 218 teams (872 individuals).

Table 8 Summary of Main Empirical Results & Consistency with Theoretical Perspectives

	Result	Diverse Recombination Perspective	Division of Specialized Labor Theory	Social Processes Perspective
1	Intermediate levels of overall diversity led to highest innovation performance; controlling for innovation quality, novelty increases with greater diversity (Section 4)	✓	✓	✓
2	Adding team members with knowledge that is most related to the problem (Business, Computer Science, Design, Engineering) boosts innovation performance; more intellectually distant knowledge (Education, Health and Nursing, Humanities, Law, Natural Sciences) has no such direct effect (Section 5.2)		✓	
3	Subgroups (of those with related knowledge) are critical to innovation performance; adding just one team member with related knowledge has no effect; subgroups are more important than the specific source of knowledge (Section.iii)			✓
4	"Balanced" teams with 2 of one related knowledge and 2 from another outperform all other configurations; the effect is more than double the effect of one subgroup of 2; similar effects are found in gender-balanced teams (Section 5.3.iii)			✓
5	The effect of subgroups on *Quality* largely (but not entirely) comes from teams with subgroups more effectively coordinating and mobilizing (successfully making a completed submission) (Section 5.3.iii)			✓
6	Although subgroups help *Novelty* from effective coordinating and mobilizing, having multiple members of same knowledge reduces Novelty all else being equal (Section 5.3.iii)	✓	✓	✓
7	Diversity per se leads to higher innovation *Quality* and even higher *Novelty* (Section 5.3.iv)	✓		

Appendix A: Experimental and Causal Studies of Team Diversity and Outcomes

Study[20]	Sample Size	Nature of Subjects	Dimensions of Diversity Studied	Primary Relationships Reported
(Aggarwal and Woolley 2019)	Study 1: 70 teams of 2 Study 2: 64 teams of 2	University students lab subjects	Cognitive styles (object and spatial visualization)	Study 1: Spatial visualization correlated with more process-focused teams. Study 2: Cognitive style heterogeneity negatively correlated with strategic consensus and increased errors.
(Aggarwal et al. 2019)	98 class project teams	MBA students	Cognitive styles	Inverted-U correlation between the sum of standard deviations in cognitive survey responses and average task performance.
(Hamilton and Nickerson 2003)	23 teams	Workers at manufacturing plants	Ability levels	Mix of high-ability and low-ability workers performed better than homogenous teams, suggesting benefits from diversity in skill levels.
(Hoogendoorn, Parker, and Van Praag 2017)	49 teams class project teams	Business students at Amsterdam College	Cognitive abilities	Significant relationship between cognitive differences and team performance, but not with the average cognitive score.
(Lyons 2017)	162 teams of 2	Contractors from India, Pakistan, and Bangladesh	Computer languages knowledge (JavaScript and PHP)	Teams improved outcomes for contractors in nationally homogeneous teams but worsened outcomes for contractors in nationally diverse teams.
(Marx, Pons, and Suri 2021)	30 teams of 2	College graduates in Nairobi	Educational levels (undergrad vs. grad)	Teams with different education levels completed more home visits and spent more time at each visit for voter registration.

[20] NB. Not all studies reported as experiments are causal, depending on the nature of relationships and measures they report.

(Jeffrey T. Polzer, Milton, and Swarm Jr 2002)	83 teams of about 5 on average (assigned to maximize diversity, rather than random assignment)	First-year MBA students	U.S. citizenship, race, sex, MBA concentration, and previous job function ; degrees prior to entering MBA ((business, engineering, liberal arts, science, other).	Diversity tended to improve creative task performance in groups with high "interpersonal congruence" (alignment of perceptions within the group, e.g., "I think I am creative, does the group also think I am creative") while it undermined performance of groups with low "interpersonal congruence."
(Rosendahl Huber et al. 2020)	112 teams	Adolescent children taking an entrepreneurship training class	Math and verbal skills balance	Teams with balanced math and verbal skills performed better than those with a mix of stronger math or verbal skills.
(Woolley et al. 2010)	40 teams of 4-5	Lab experiment participants	Loading multiple variables onto a component factor analysis, with one factor referred to as "C"	Task performance across multiple tasks is correlated with the mapping of multiple survey variables to a primary component analysis factor (referred to here as "C"), and once accounting for the projection onto this best fitted factor, performance is not strongly correlated wtih either the average or maximum individual intelligence (based on pre-experiment test) of the team members.
(Woolley et al. 2008)	41 teams of 4	Boston-area students and residents	Subjects with high scores in verbal memory vs face recognition abilities	Teams assigned with subjects with high verbal memory or facial recognition (based on pre-experimental testing) required collaborative planning (a requirement discuss which team members would assess which type of evidence) to perform most effectively.

(Hansen, Owan, and Pan 2006)	102 student groups	Undergraduate management class students	Gender, age	Male-dominant groups performed worse in both group work and individual exams compared to female-dominant and mixed-gender groups. Diversity in age and gender correlated with higher exam scores. Democratic contract had a positive effect on performance.
(Van Knippenberg, Haslam, and Platow 2007)	220 business students in teams of different sizes	Participants of a cross-sectional survey and a laboratory experiment	Gender diversity and diversity beliefs	Work group diversity and group identification are more positively related the more individuals believe in the value of diversity

Appendix B: Product Development Platform

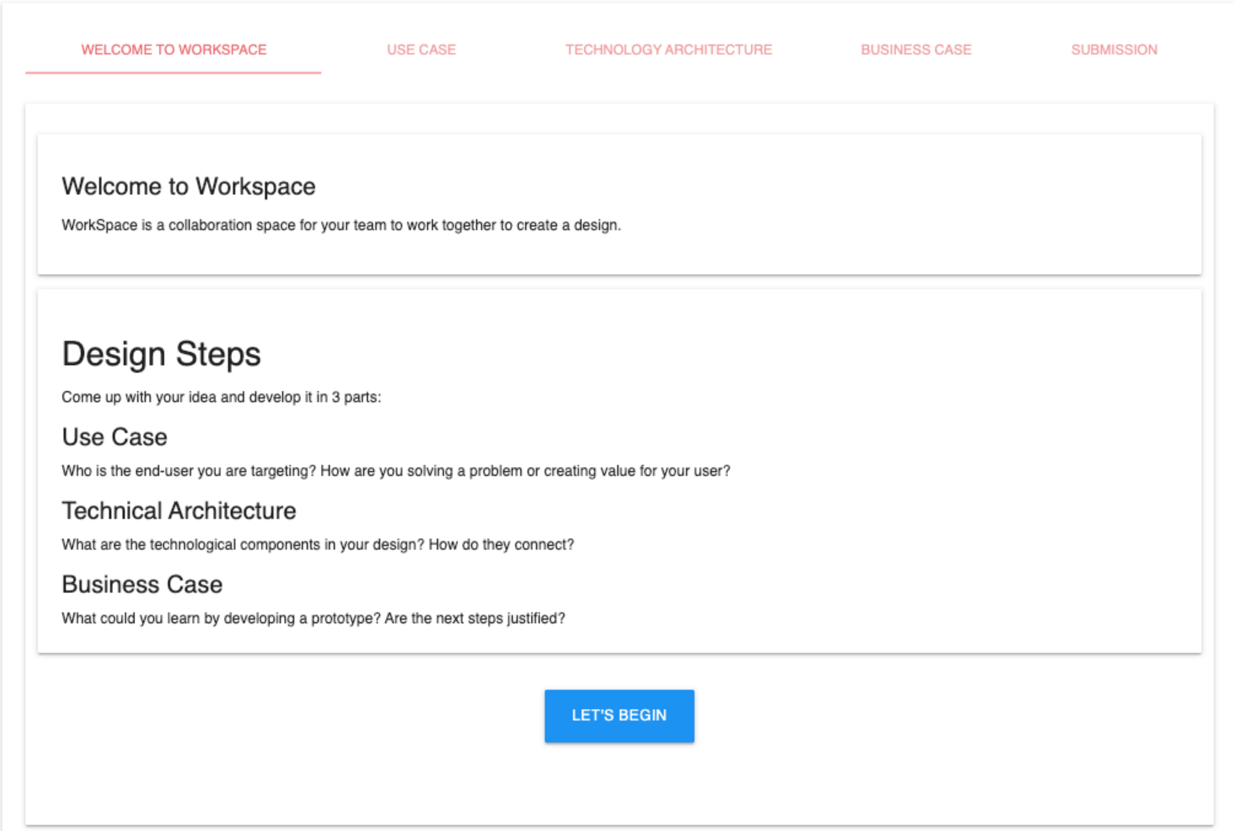

Fig. B1 Description of IoT Project Work: Design of Use Case, Technology Architecture, and Business Case for Prototype

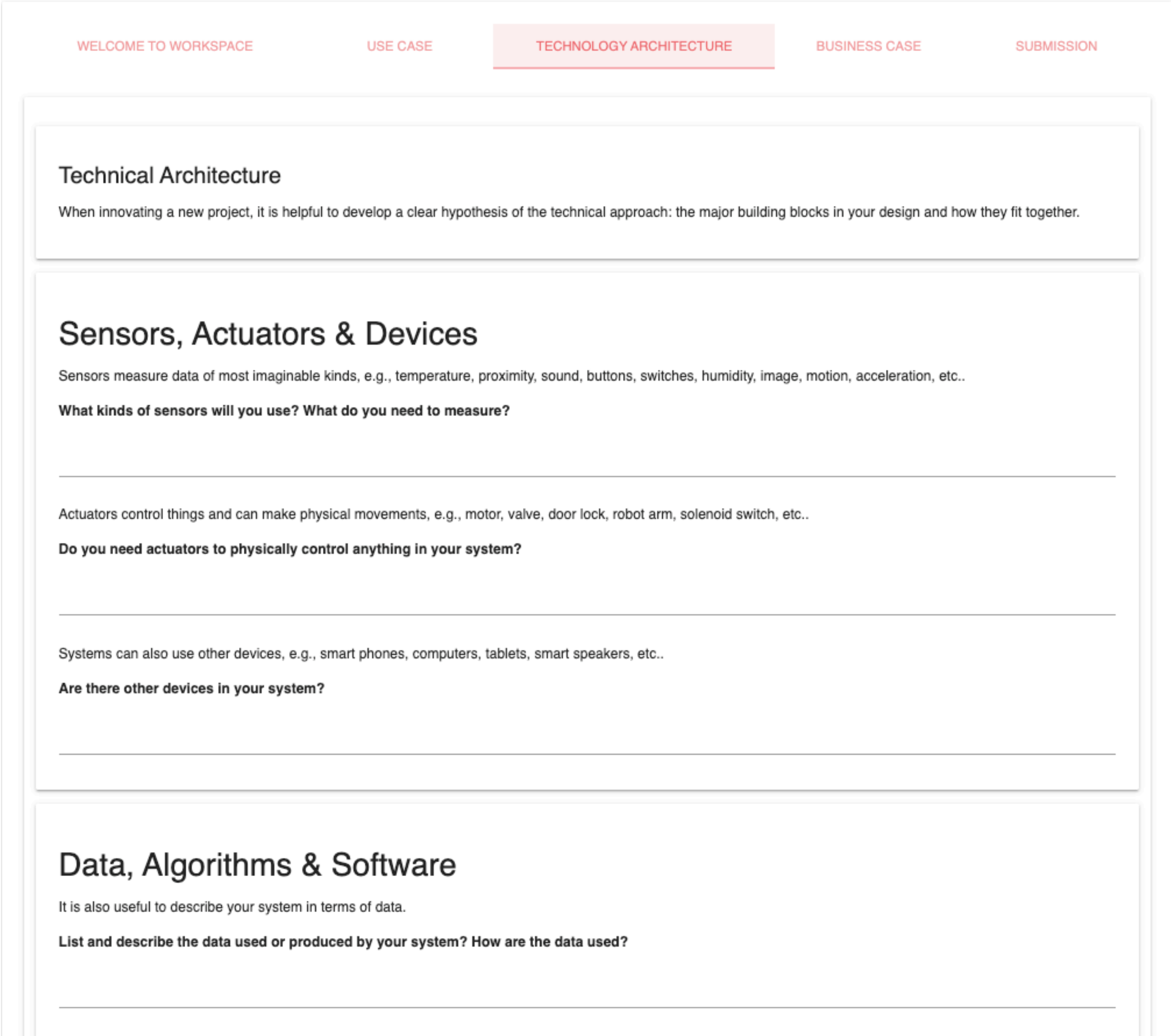

Fig. B2 Sample Section of Technology Architecture Design Steps

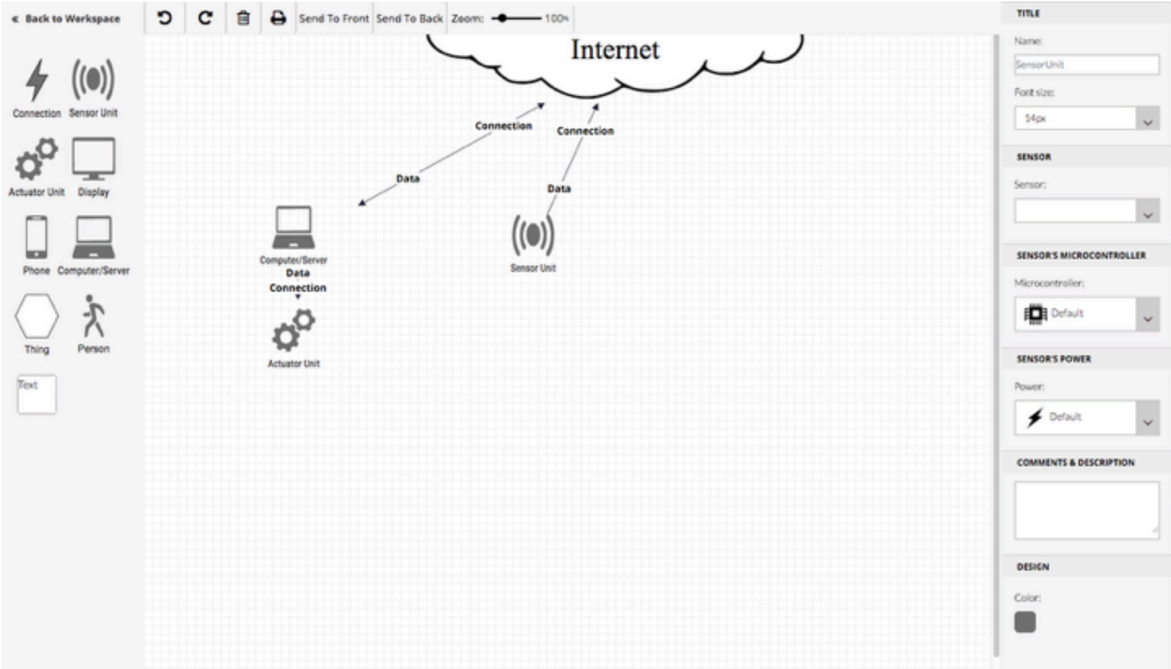

Fig. B3 Sample Technology Design Drag-and-Drop Design Tool

Appendix C: Additional Analysis & Regression Tables

i. Ineffectiveness of Simple Mixing from Related Fields

Given that adding team members from each of Business, Computer Science, Design, and Engineering leads to superior performance, we might think that randomly drawing from this pool of related knowledge to assemble teams would lead to superior performance. This is not the case. Here, we regress overall innovation *Quality* on dummies counting whether there are 1, 2, 3, or 4 team members from some related field. Thus, this regression effectively estimates the effect of a manager simply randomly drawing from the pool of related knowledge to assemble teams.

As reported in model (1) of Table 9**,** even teams with 4 team members from related fields, randomly drawn, do not lead to a significant positive effect on innovation quality, and having fewer can leads to negative effects.[21] Therefore, if there are benefits of combining knowledge, it may take more specific combinations. We similarly find no evidence of positive effects when regressing *Novelty* on these dummies, whether unconditional of *Quality* as in model (2), or conditional on *Quality* as in model (3). Therefore, if there are benefits of certain combinations of skills, they must take a more particular form than just randomly drawing from the pool of individuals with related knowledge. A similar lack of effects is found in the case of regressions of *Novelty* (model 2), and Novelty conditional on *Quality* (model 3).

[21] Although it is plausible that the negative effect is caused by including certain counts or combinations of related types, there is a more straightforward explanation for the negative signs. For example, where the number of related types is 1, this means that *all* others in the team are not randomly drawn, but rather they are necessarily from unrelated fields.

Table 9 No Evidence that Simply Randomly Drawing from Related Knowledge Types Leads to Performance

Dep. Var.:	*Quality*	*Novelty*	
Model:	(1)	(2)	(3)
Team Member Counts by Knowledge Fields			
"Related" Fields - Proximate to Problem:			
Related = 1	-2.50*	-0.97	-0.12
	(1.40)	(1.43)	(0.15)
Related = 2	-1.93	-1.58**	0.00
	(1.27)	(0.67)	(0.09)
Related = 3	-1.68	-1.65	-0.05
	(1.24)	(1.26)	(0.12)
Related = 4	-0.28	-0.34	-0.08
	(1.28)	(1.30)	(0.13)
Health & Nursing	0.27	0.37	0.11
	(0.54)	(0.54)	(0.09)
Humanities	-0.28	-0.24	0.03
	(0.39)	(0.37)	(0.05)
Law	0.01	-0.10	-0.11
	(0.87)	(0.82)	(0.09)
Sciences	-0.12	-0.12	0.00
	(0.35)	(0.34)	(0.06)
Grad Degree	0.18	0.22	0.04*
	(0.17)	(0.16)	(0.02)
Controls			
Gender	Y	Y	Y
Age	Y	Y	Y
Race	Y	Y	Y
Quality			0.95***
			(0.01)
Constant	0.97	.79	-.13
	(2.79)	(2.64)	(0.35)
Adjusted-R^2	0.03	0.11	0.98

Notes. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. OLS model coefficient estimates; robust standard errors in parentheses; gender controls = No. of women; age controls = average and range of ages; race controls = counts (see Section 3.4 discussion). No. obs. = 218 teams (872 individuals).

ii. Ineffectiveness of Simple Linear Complementarities (Interactions)

We also might investigate whether more specific related knowledge combinations are effective, here examining the most straightforward possible linear interactions, i.e., *Business* × *Design, Business* × *Engineering, Computer Science* × *Design, Computer Science* × *Engineering, Design* × *Engineering*. Models (1) to (6) of Table 10 reports models in which each of these interactions are added individually, as well as all at once. None of these simple interaction terms is statistically different from zero. (If control variables are dropped, the interaction between *Computer Science* × *Engineering* becomes marginally statistically significant.) This is also true when regressing *Novelty* on these interactions, as in model (8), or doing so while also conditioning on overall innovation *Quality*, as in model (9).

Table 10 No Evidence of Simple Linear Complementarities of Knowledge (Interactions)

Dep. Var.:	*Quality*							*Novelty*	
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Interactions between Different Fields									
Business × Computer Sci	0.13						0.67	0.68	0.04
	(0.45)						(0.59)	(0.55)	(0.09)
Business × Design		0.27					0.41	0.51	0.12
		(0.51)					(0.77)	(0.74)	(0.12)
Business × Engineering			-0.10				0.09	0.06	-0.03
			(0.35)				(0.40)	(0.38)	(0.05)
Computer Sci × Design				-0.07			0.58	0.57	0.02
				(0.92)			(0.96)	(0.94)	(0.13)
Computer Sci × Engineering					0.80*		1.17*	1.16*	0.05
					(0.46)		(0.61)	(0.60)	(0.07)
Design × Engineering						-0.25	0.17	0.28	0.12
						(0.56)	(0.75)	(0.72)	(0.08)
Team Member Counts by Knowledge Fields									
Counts by Knowledge Fields	Y	Y	Y	Y	Y	Y	Y	Y	Y
Grad Degree	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls									
Gender	Y	Y	Y	Y	Y	Y	Y	Y	Y
Age	Y	Y	Y	Y	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y	Y	Y	Y	Y
Quality									0.95***
									(0.02)
Constant	0.45	0.40	0.19	0.40	0.78	0.38	1.43	1.32	-0.04
	(2.82)	(2.79)	(2.94)	(2.80)	(2.81)	(2.79)	(3.08)	(2.91)	(0.36)
Adjusted-R^2	0.16	0.16	0.16	0.16	0.18	0.17	0.16	0.17	0.98

Notes. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. OLS model coefficient estimates; robust standard errors in parentheses; gender controls = No. of women; age controls = average and range of ages; race controls = counts (see Section 3.4 discussion). No. obs. = 218 teams (872 individuals).