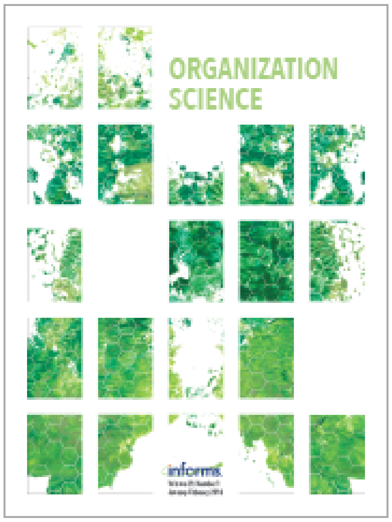


Organization Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Extemporaneous Coordination in Specialist Teams: The Familiarity Complementarity

Kenny Ching, Enrico Forti, Evan Rawley

ORGANIZATION SCIENCE
Articles in Advance, pp. 1–17
ISSN 1047-7039 (print), ISSN 1526-5455 (online)
http://pubsonline.informs.org/journal/orsc

Extemporaneous Coordination in Specialist Teams: The Familiarity Complementarity

Kenny Ching,[a] **Enrico Forti,**[b,c,d] **Evan Rawley**[e]

[a] Foisie School of Business, Worcester Polytechnic Institute, Worcester, Massachusetts 01609; [b] UCL School of Management, University College London, London, E14 5AB, United Kingdom; [c] O'Malley School of Business, Manhattan College, New York, New York 10471; [d] Columbia University, Columbia Business School, New York, New York 10027; [e] Carlson School of Management, University of Minnesota, Minneapolis, Minnesota 55455

Contact: kching@wpi.edu, https://orcid.org/0000-0002-8049-9133 (KC); e.forti@ucl.ac.uk, https://orcid.org/0000-0003-4289-4244 (EF); erawley@umn.edu, https://orcid.org/0000-0003-3916-3942 (ER)

Received: November 3, 2018
Revised: September 20, 2019; December 13, 2019
Accepted: January 3, 2020
Published Online in Articles in Advance: December 7, 2020

https://doi.org/10.1287/orsc.2020.1376

Abstract. Team production is ubiquitous in the economy, but managing teams effectively remains a challenge for many organizations. This paper studies how familiarity among teammates influences the performance of specialist teams, relative to nonspecialist teams. Applying theories of team production to contexts where team members coordinate interdependent activities extemporaneously, we develop predictions about factors that shift the marginal returns to specialization along two dimensions of familiarity: social familiarity and functional familiarity. We test our hypotheses in the context of Defence of the Ancients 2 (DOTA2), a major e-sports game where, in some formats, players are exogenously assigned to five-person teams. After analyzing nearly 6.5 million matches, we find that specialist teams are relatively more successful when members are more socially and functionally familiar with one another. The results suggest that the plug-and-play perspective on specialist teams is incomplete; rather, specialization and familiarity are complements in dynamic environments where team members coordinate extemporaneously.

Funded: Financial support from the UCL School of Management and Worcester Polytechnic Institute is gratefully acknowledged.
Supplemental Material: The online companion is available at https://doi.org/10.1287/orsc.2020.1376.

Keywords: organization • familiarity • specialists • teams • performance • e-sports

Introduction

As the growth of project work alters the way firms manage human capital, organizations increasingly mobilize temporary teams to tackle problems. Thus, modern economic activity is often performed by teams of coworkers who contribute expertise to complete a task, without expectation of stable collaboration or extensive hierarchical oversight (Cattani et al. 2011, Barley et al. 2017, Kuhn and Maleki 2017). To be successful in dynamic environments with highly interdependent tasks, good project teams must coordinate extensively (Stan and Puranam 2017).[1] But, whereas the importance of coordination has been recognized in the literature on teams (e.g., Rico et al. 2008, Gardner et al. 2012, Srikanth et al. 2016), it is unclear how the need for meaningful extemporaneous coordination (rapid mutual adaptation to unexpected circumstances, in a context with nondecomposable, but highly interdependent tasks) alters the optimal organization of teams.[2] This paper advances our understanding of how to design and manage teams by studying how team composition and team familiarity jointly influence performance, in contexts where extemporaneous coordination is important. Specifically, we analyze how familiarity (being well acquainted with one's teammates) influences the performance of specialist teams—where specialists are defined as those who have accumulated experience within a narrow range of tasks—in environments where the optimal allocation of tasks to team members is dynamic and unpredictable.

Positive returns to specialization have been widely reported in many settings. Economic accounts note that specialists are more productive than nonspecialists and require less training during onboarding (e.g., Gibbons and Waldman 2004, Staats and Gino 2012), whereas sociological studies find that specialists are easier to categorize and, absent status signals, outperform nonspecialists in social evaluation processes (e.g., Ferguson and Hasan 2013, Zuckerman 2017). But, recent studies show that teams of specialists sometimes experience coordination problems (Reagans et al. 2005, 2016), implying a tension between deploying and coordinating specialist teams in certain dynamic environments (Ben-Menahem et al. 2016). One clear implication of the research on the limitations of specialized teams is that the returns to specialization will vary as the context within which a team operates changes. And, yet, even though project teams have become ubiquitous, there is little research on the specific contingencies that systematically influence the returns to specialization in different contexts.

In this paper, we shed some light on how specialized project teams operate in dynamic interdependent environments, formulating and testing a simple, but novel, theory of how familiarity influences the relative efficacy of specialist teams. Specifically, we propose that at least two dimensions of familiarity influence how effectively specialist teams coordinate their actions: social familiarity and functional familiarity. Social familiarity, arising from repeated interactions between team members, gives teammates insight into one another's preferences, styles, and idiosyncrasies (Deming 2017). Functional familiarity, arising from performing similar roles independently, gives teammates insight into the subtleties, difficulties, and opportunities associated with performing a particular role.

Though conceptually distinct, both dimensions of familiarity impact coordination in similar ways. Specialists, by definition, have less experience performing the other tasks required of the team, and, therefore, have a more limited understanding of how to coordinate with others on the team when the optimal nexus of tasks shifts. For a team of specialists, familiarity acts as a substitute for members' lack of knowledge of how one another's tasks are performed. By contrast, nonspecialists have a broader set of experiences to draw upon, which facilitates mutual understanding and coordinated adaptation. Nonspecialists, therefore, need rely on familiarity less to effectively coordinate. In highly structured contexts, specialists can coordinate effectively by designing modular tasks—they "plug and play" (Okhuysen and Bechky 2009). But, in environments where teammates must jointly adapt to unexpected circumstances, specialist teammates who understand one another better, socially and functionally, will be marginally more effective than nonspecialist teams with the same level of familiarity.

The empirical tests of the theory exploit a large and rich panel data set on teams in a setting where extemporaneous coordination is fundamental to team performance: e-sports. Specifically, we study teams competing in Defence of the Ancients 2 (DOTA2), a popular high-stakes competitive strategy video game. Although e-sports is a relatively new phenomenon, it has exploded in popularity, attracting close to $1.5 billion in venture capital in 2018, with industry analysts predicting revenues in excess of $1 billion from sponsorships, media rights, and ticket sales in 2019 (Merwin et al. 2018). Scholars, recognizing the opportunity to use e-sports as a laboratory for testing organizational theories, have become increasingly interested in the phenomenon too (Waguespack et al. 2018, Clement 2019). Importantly, the context allows us to exploit exogenous assignment of players to teams (in a large subset of the matches), providing millions of organizational design experiments to analyze. Moreover, DOTA2 teams are qualitatively similar to project teams in traditional businesses, particularly those operating in dynamic interdependent environments such as consulting teams, pharmaceutical research teams, and teams of software developers, suggesting that the insights from this research should apply to a wide range of team production contexts.

The results demonstrate that both social and functional familiarity are complements to specialist teams, suggesting that specialist teams do more than plug and play, they also learn to harmonize their actions by understanding their teammates better. More broadly, the results speak to the importance of systems of team characteristics, which can shift the marginal returns to team production. Specifically, we show that it isn't familiarity or specialization alone that drives team performance in dynamic interdependent settings, but also the interplay between them.

Theory and Literature

At least since the seminal work by Smith (1776) on the division of labor, and the studies of task allocation by Taylor (1911), scholars have endeavored to understand the connection between specialization and performance. Many scholars have noted that specialization, defined as focused experience at a task, tends to yield higher productivity at that task (e.g., Becker 1962, Schilling et al. 2003). Moreover, teams of specialists also tend to outperform (e.g., Valentine and Edmondson 2015). However, organizational scholars have observed that when teams operate in settings characterized by highly interdependent and uncertain tasks, specialization becomes less valuable (e.g., Reagans et al. 2016). As the optimal division of tasks becomes less stable, coordination challenges increase (Stan and Puranam 2017). For specialist teams in such contexts, a lack of "trans-specialist understanding" (Postrel 2002, p. 303)—mutual understanding among specialist teammates—makes coordinated adaptation even more challenging.

Considering that organizations increasingly face environments characterized by interdependent nonroutine tasks that "cannot be specified fully through design" (Bechky and Chung 2017, p. 2)—precisely the kind of setting where specialists may struggle to coordinate effectively—it seems important to understand how organizational design might improve specialist teams' performance in such settings. Thus, this paper addresses a gap in the literature by studying performance of specialist teams in contexts where mutual adaptation to unforeseen circumstances improves coordination. Given an environment where extemporaneous coordination is important, we focus, in particular, on a characteristic of teams that seems particularly likely to improve coordination—how familiar team members are with one another.

Extemporaneous coordination differs from explicit coordination, the ability of a team to discuss and deliberate on plans, responsibilities, roles, and tasks ex ante. Indeed, extemporaneous coordination is a dynamic adjustment process where team members mutually adapt their behavior in response to surprises that change the optimal interdependencies between tasks in unpredictable ways. To evaluate the impact of familiarity on specialist teams that coordinate extemporaneously, we study familiarity along two distinct dimensions: social familiarity, arising from repeated interactions between team members, and functional familiarity, arising from team members having performed similar roles independently.

Familiarity is an important concept in its own right,[3] but to the extent that it facilitates the harmonization of unpredictable interdependent tasks, it can be a potent shifter of the marginal returns to team composition. Indeed, empirical evidence shows that teams characterized by low levels of social familiarity between members (e.g., new teams and teams that experience unplanned membership changes) operate differently from teams with higher levels of social familiarity. For example, unfamiliar teams tend to organize functionally based on the skills of new members (Berman et al. 2002, Bechky 2006), exert more effort negotiating task allocation (Skilton and Dooley 2010), and may be less efficient cognitively (Lewis et al. 2007). Relatedly, other research has shown that teams with better social skills coordinate more effectively (Deming 2017). The coordination benefits of social familiarity become particularly relevant when team members face unknown and unpredictable interdependencies that require extemporaneous coordination (Barley et al. 2017, Bechky and Chung 2017).

Existing studies broadly support the idea that some forms of familiarity lead to improved team coordination and performance, suggesting that familiarity facilitates a shared understanding of teammates' unique and idiosyncratic behaviors under different environmental conditions (Lewis et al. 2005, Ren and Argote 2011). However, little is known about how familiarity influences specialist teams relative to nonspecialist teams.

Both social and functional familiarity affect the development of cognitive, social, and organizational systems based on trust (Jarvenpaa and Leidner 1999, Majchrzak et al. 2007) and knowledge about who knows what in a group (Lewis et al. 2005, Argote and Miron-Spektor 2011), allowing individuals to communicate and coordinate effectively within the team (Ren and Argote 2011). Social familiarity and functional familiarity are, however, conceptually and practically distinct, reflecting, respectively, the frequency of interaction between teammates and the shared roles that teammates have performed independently in the past.

Although familiarity sometimes breeds contempt, one might expect that social familiarity would usually create meaningful advantages for teams, particularly when extemporaneous coordination is important. Teammates who know one another well should be better able to anticipate their teammates' actions, and therefore be more prepared to adapt quickly to their teammates behaviors (Rico et al. 2008, Ren and Argote 2011). Functional familiarity, however, can clearly be a double-edged sword. Having a similar functional background may facilitate mutual understanding and communication between teammates, given their shared understanding of how one should behave under different contextual conditions of the work environment (Luciano et al. 2018). However, ceteris paribus, functional familiarity implies redundancy, which means teams of a fixed size must sacrifice breadth of knowledge in return for functional familiarity. Thus, though the main effect of social familiarity on performance should be expected to be positive, the main effect of functional familiarity is uncertain.

In this paper, we are specifically interested in the marginal effects of social familiarity and functional familiarity on specialist teams relative to nonspecialist teams. On the one hand, familiarity might have no marginal effect on specialist teams—familiarity effects could be the same regardless of team composition. For example, if specialists operate in a context where key tasks can be substantially decomposed, and interdependencies effectively managed through a clear division of labor, then thoughtful job design will be sufficient to achieve effective coordination among teammates, and familiarity will not meaningfully substitute for a lack of shared task knowledge. And, indeed, the literature on role-based coordination in teams finds that, in some contexts, the focused identity of specialists aids integration and coordination in groups, even in the absence of social familiarity between members (e.g., Bechky and Chung 2017). In such contexts, specialization enables plug-and-play coordination between strangers—modular interdependency enabled by standards—allowing the successful completion of interdependent tasks (Bechky 2006).

On the other hand, plug and play can break down when the optimal interdependency between teammates changes in unexpected ways—precisely when the need for extemporaneous coordination is high—which may explain why teams composed of specialists can experience problems coordinating activities.

For example, recent work on teams, such as film crews, SWAT teams (Bechky and Okhuysen 2011), and police detectives (Schakel et al. 2016), suggests that specialists may encounter difficulties coordinating effectively when exposed to surprises that influence the optimal nexus of activities. However, familiarity can offset features of unstable task environments that impose limits on specialists' ability to coordinate extemporaneously. As members become more familiar with one another, teams tend to develop cognitive systems that improve an individual's ability to correctly predict which actions other team members will take and at what time (Lewis et al. 2005, Puranam et al. 2012). Thus, by facilitating deeper mutual understanding, familiarity improves the ability of specialist teams to coordinate extemporaneously.

Because familiarity ameliorates coordination problems in unpredictable environments, it should be uniquely valuable to specialist teams. Comparatively, nonspecialist teams, already rich in mutual task knowledge, need rely on familiarity less to aid in coordination. Thus, we propose the following hypotheses.

Hypothesis 1. (a) *In contexts requiring extensive extemporaneous coordination, increasing social familiarity will improve the relative performance of specialist teams.* (b) *In contexts requiring extensive extemporaneous coordination, increasing functional familiarity will improve the relative performance of specialist teams.*

Though intuitive, the hypotheses are by no means obvious. The magnitudes of the marginal effects of the two distinct dimensions of familiarity on specialist teams, relative to nonspecialist teams, depend upon whether specialist teams are effective solely because they plug and play well together—in which case we would not find support for the hypotheses—or whether specialists overcome their difficulties managing extemporaneous coordination by becoming more familiar with how to play well together. Taken together, our hypotheses makes a simple, broad prediction—that familiarity and team specialization are complements when extemporaneous coordination is important—a prediction we take to the data, after describing the empirical context in more detail next.

Institutional Context: DOTA2

Overview

We test our hypotheses using data from DOTA2, a popular competitive strategy video game where teams of five players coordinate extensively as they play head-to-head against another team in a zero-sum game.[4] DOTA2 is one of the most popular e-sports games in the world, with over a million concurrent players, and is widely regarded as one of the most iconic and storied team-based electronic games (Merwin et al. 2018). It has a global, vibrant, and well-developed competitive scene, the DOTA Pro Circuit, where professional players compete for cash prizes comparable to prizes paid in traditional sporting events such as golf and tennis. For example, one DOTA2 tournament (The International) offered a total prize pool of more than $33 million in 2019.

DOTA2 is a particularly attractive setting for studying how familiarity influences the returns to specialization for several reasons. DOTA2 teams closely resemble project teams in professional settings. Just as with DOTA2 teams, new product development teams, consulting teams, and teams of application developers are often structured to be temporary, relatively flat organizations, assembled for the purpose of achieving a well-defined outcome in environmentally turbulent contexts. Just as with DOTA2, members of temporary project teams in more traditional professional settings bring heterogeneous skill sets to the group that must be harmonized and coordinated in the face of environmental changes. And, just as with DOTA2, the effectiveness of temporary teams of specialists in other professional settings will be influenced by how socially and functionally familiar teammates are with one another. Thus, DOTA2 offers a valuable laboratory in which one can study the marginal effect of familiarity on team composition.

Additionally, the DOTA2 laboratory allows one to reliably measure the key conceptual constructs from the theory: specialization, team composition, social familiarity, and functional familiarity. The gaming platform records detailed statistics on DOTA2 matches, including data on individual players who are tracked with a unique identifier that is consistent across matches. Using this data, one can create detailed career histories of players, which can be characterized in organizationally meaningful ways.

Forming Teams and Playing the Game

Most DOTA2 matches are organized via a matchmaking algorithm that exogenously assigns individual players to teams with the objective of making matches competitive (that is, DOTA2 randomizes on expected outcomes based on players' track records) providing experimental variation in the assignment of players to teams that we exploit to identify the causal impact of the marginal effect of familiarity on team composition. Although matchmaking is not completely random in all games, we verify that players were effectively exogenously assigned to teams in a subset of the data (i.e., the subset we use).[5]

During match play, each player on a team controls exactly one of the 112 game characters, known as heroes, who have unique characteristics and abilities. For example, certain heroes are designed for shorter and focused activity whereas other heroes are more suitable for long and drawn out endeavors. Although game rules change in ways that make the value of heroes fluctuate over time, among the 112 heroes are two broad types that are particularly germane for our study.[6] Certain heroes are associated with *carry* positions—offensive players who conduct the team to victory late in the game. Other heroes are associated with *support* positions—players who help the carry players prepare for their attempt at victory. Carry and support players are differentiated, but perform complementary roles, which are highly analogous to roles played by teammates in traditional sports, and similar in spirit to many traditional business settings where roles are differentiated.[7] We designate players who typically play either carry or support *specialists*, and players who often play both carry and support *generalists*.

In the matches we study, where teams are exogenously assigned, extemporaneous coordination is manifest less by one player formally leading the team, and more through a range of interdependent activities taken in consideration of other players' roles, abilities, and actions, as well as opponents' behavior. Many of these activities are actions that advance the cause of the team in subtle ways that may not even be well understood by all members of the team at the time the action is taken. Thus, DOTA2 is a context where one should expect that familiarity would improve extemporaneous coordination among teammates.

Selecting Roles

Following exogenous assignment into a team, players have a few minutes to choose the hero they want to play in the focal match in consultation with their teammates. Since certain configurations of heroes may dominate others, nonspecialists, and particularly generalists, add value to their teams by selecting into heroes that fit well with the needs of the team. Specialists too have a choice over which hero to play, though they are more constrained given their more limited breadth of experience. Indeed, specialists play off-character in only about 7% of all matches in our sample.

Although hero selection decisions are clearly endogenous—they happen after exogenous assignment of players to teams, but before the play of the game begins—an individual's choice of which hero to play does not influence the measurement our key explanatory variables. Specialization and familiarity are measured based on player histories, not on choices made in the focal game. Since the theory advances the idea that familiarity facilitates mutual understanding, regardless of which heroes teammates choose to play, endogenous role selection is not a conceptual concern either. Nevertheless, endogenous hero selection does represent a challenge to inferring from the data whether our proposed mechanism—extemporaneous coordination—is at work. We acknowledge that we cannot completely rule out the possibility that familiarity influences performance both by improving extemporaneous coordination and ex ante hero selection decisions, since players who know each other well may find it easier to agree on the "right" hero choices. However, as we discuss in more detail in the Discussion section, the full set of results we present suggests there is compelling evidence in favor of the extemporaneous coordination mechanism.

Match Types

DOTA2 organizes three kinds of matches: ranked matches, unranked matches, and professional matches. Ranked matches contribute to a rating system called *match making rating*, whereas unranked matches do not. In most ranked matches, all 10 players are exogenously assigned to teams, though there are some exceptions that we describe later. Importantly, we use all matches to measure specialization and familiarity, but only ranked matches for our main analyses, (though the results are similar when we include unranked matches too). The third type of match (professional matches, where all teams are formed endogenously) are used to identify the key players of interest, but are not included in the test sample.

DOTA2 hosts three types of ranked matches: solo player, team play, and party play. In ranked matches (the game format we focus on) the gaming algorithm designed by DOTA2's publisher uses individual histories, much like chess ratings, to level teams so that each team is equally likely to win the match. Importantly, though, the matching algorithm does not consider specialization or familiarity.[8] Most ranked matches are solo player games, where individual players choose when to join the gaming platform, and are then automatically sorted into tiers contemporaneously, based on their prior performance.[9]

Data and Sample

Using Dotabuff.com, a publicly available website chronicling the professional DOTA2 scene, we identified 4,272 unique players who ever participated to at least one professional DOTA2 match—the serious players. We then downloaded data on all matches ever played by each professional player using the DOTA2 WebAPI service,[10] allowing us to track the full careers of the serious players, including all the matches where teams are exogenously assigned, from the inception of the game in 2011 through to the end of 2016.[11] The full data set of matches—professional matches, unranked matches, and ranked public matches—including at least one professional player consists of 306,949 individual players participating in nearly 9.2 million matches over a five-year period. We use the full sample to compute all of our explanatory variables, whereas our main results are estimated on the subset of 6,444,502 ranked public matches with at least three nonanonymous players per team.[12]

Variables and Measures

Our dependent variable is *Victory*, an indicator set to 1 if the focal team wins the match, and 0 otherwise. Although victory is clearly important, alternative outcome measures are more ambiguous in high-level DOTA2 matches, as victory is often dependent on a complex set of intermediate steps, and is rarely secured until late in a game. Moreover, using victory to measure the outcome of team efforts follows common practice among scholars studying the organizational performance of traditional sports teams (Smith and Hou 2015, Fonti and Maoret 2016, Stuart 2017).

The key explanatory variables in this study are the interactions between a team composition variable, *Specialist Team*, and two team familiarity variables *Social Familiarity* and *Functional Familiarity*. To calculate *Specialist Team*, we first define a player's level of specialization, *Individual Specialization*, based on the diversity of the player's past experience as a carry and support player.

Following the existing literature, we measure individual specialization using a concentration index (e.g., Zuckerman et al. 2003, Narayanan et al. 2009, Ferguson and Hasan 2013). The measure captures the idea that at some point a player's experiences become so concentrated that the player should be considered a specialist. For example, players that usually play either the carry or support role, but not both, will have highly concentrated role experience, and will be classified as specialists. Specifically, *Individual Specialization* is computed by a Herfindahl-Hirschman index (HHI), compared with a set of players with the exact same amount of cumulative experience: $\mathit{Individual\ Specialization} = \Sigma_{\mathrm{r}} (S_{irt}/S_{it})^2$, where S_{irt} is the total number of times an individual i has played in role r by match t of i's tenure, and S_{it} is the total number of matches i has played through match t. Players are compared with others with the same level of experience to ensure experience effects do not contaminate the specialization measure. Larger scores, those closer to one, are associated with increased specialization, whereas smaller scores, those closer to zero, capture less specialization (more generalism).

Since we ultimately want to create a team-level specialization construct, discretizing individual specialization is a natural modeling choice: conceptually, we want to know if teams are comprised of a set of specialists, not whether the average specialization score of the team is high. Thus, consistent with the literature on specialization, we define a player as a specialist if the player is above a certain percentile of the *Individual Specialization* distribution (e.g., Teodoridis 2018, Teodoridis et al. 2018). Since the HHI index is computed among players with the same depth of experience, the measure is also consistent with Lazear (2004). Because there is no natural absolute measure of specialization, we conservatively use the 75th percentile of the *Individual Specialization* distribution as the cut-off for characterizing a player as a specialist, though the results are robust to using other cut-off values (e.g., the median) or a continuous measure of specialization (i.e., the mean of the team's *Individual Specialization*).

In our main specification, we define *Specialist Team* as an indicator variable that is equal to one if the team is comprised of a majority of specialists—that is, if it has three or more individual specialists. In a series of robustness checks, we show that the results are not sensitive to other reasonable definitions of *Specialist Team* (e.g., four or five specialists). Similarly, we define a player as a generalist if the player is below the 25th percentile of the *Individual Specialization* distribution, among all players that played the same number of matches. It then follows that a *Generalist Team* is one where a team has three or more individual generalists—those in the lowest 25th percentile of the specialization distribution. All other teams are nonspecialist/nongeneralist teams—neither specialist nor generalist teams.

To measure *Social Familiarity*, as in previous studies (e.g., Reagans et al. 2005, Huckman et al. 2009), we first measure the level of individual familiarity that team members have with their teammates by counting the number of (dyadic) times they played previously with any of their other teammates. Analogously, for each team in a given match, we measure *Social Familiarity* as the mean of the teammates' individual familiarity. To facilitate interpretation, we standardize *Social Familiarity* to be mean zero and have a standard deviation equal to one.

The variable *Functional Familiarity*, computed using each player's history of hero choices, represents the extent to which, within a team, members are similar to one another in terms of their functional backgrounds. As in prior studies, we compute, for each player on each team, an average cosine similarity index (Sohn 2001, Choudhury and Haas 2018), which generates a scale-free (i.e., a measure that is independent of the number of matches one has played) similarity index of a player's portfolio of hero-specific experiences compared with that of their teammates. Specifically, individual functional familiarity is computed as the cosine of the angle between two vectors that include counts of the total number of times that each dyad of players within the team played any of the 112 heroes in the past. For each dyad of players, the values of the measure range between 0 and 1, where 0 is a perfectly dissimilar and 1 is an identical functional background. We average this measure across team members to compute *Functional Familiarity* at the team level.

Figure 1. Stylized Example Illustrating the Difference Between Specialization and Functional Familiarity

	Carry		Support	
	Hero 1	Hero 2	Hero 3	Hero 4
Player 1	✪			
Player 2			✪	✪
Player 3		✪	✪	
Player 4	✪	✪		

Notes. The figure illustrates the empirical distinction between the constructs of *Specialization* and *Functional Familiarity* with a stylized example—where a black circle with a white star represents heroes a player has played in the past. For simplicity, we assume that players can choose among four heroes instead of 112, and that teams are dyads instead of groups of five. Consider the following four players, each of which can specialize, or not, in a carry or support role. A carry role can be executed with hero 1 or hero 2, whereas a support role can be executed with either hero 3 or hero 4. In this simplified example, Player 1 and Player 2, who are both specialists, have no functional familiarity, whereas Player 1 and Player 4, both specialists, have positive functional familiarity. Player 1, a specialist, and Player 3, a nonspecialist, have no functional familiarity, whereas Player 3, a nonspecialist, and Player 4, a specialist, have positive functional familiarity. As the example illustrates, the measures of *Specialization* and *Functional Familiarity* need not be correlated.

Our measures of *Specialization* and *Functional Familiarity* exploit two levels of aggregation in hero space. *Specialization* is measured coarsely, based on a player's history playing either carry or support heroes, which effectively captures a player's style of play. By contrast, *Functional Familiarity* is measured based on a player's fine-grained history playing each particular hero, which effectively captures a player's task experience. Although the constructs are conceptually and empirically distinct (Figure 1 illustrates the empirical distinction between *Specialization* and *Functional Familiarity* with a stylized example), in principle, the measure associated with a particular construct could be reversed. Empirically, the results in this paper are not meaningfully changed if we measure specialization at a more fine-grained level, and functional familiarity more coarsely; however, within the context of DOTA2 it is more typical (i.e., among players) to refer to specialization as a style of play, rather than as a particular hero choice. We have, therefore, used the natural distinction within the institutional context to guide our particular measures.

Individual team members in a match operate under epistemic interdependence—that is, "one agent's optimal choices depend on a prediction of another agent's actions" (Puranam et al. 2012, p. 420)—within a team, and also relative to opposing teams' actions. In other words, the competitive nature of our empirical context implies that team members on opposing teams attempt to predict and react to each other's actions. Since our theory of specialization and familiarity should apply to a rival team, just as it applies to a focal team, including opponent team variables in our regressions both controls for rival team effects and allows us to put our theory to a further test.

Descriptive Statistics

Table 1 contains descriptive statistics and correlations for the key variables used in the regressions. Since we randomly select one team from each match to include in our regressions (to avoid double counting matches), along with information on their opponents, we can evaluate whether DOTA2's matchmaking algorithm truly exogenously assigns teams by comparing the similarity of the means of the covariates between focal and opponent teams. From Table 1, one can compute the differences in means (opponent team – focal team) and t-statistics on the differences [in square brackets], which are *Specialist Team* 0.00 [0.00], *Generalist Team* 0.00 [0.00], *Functional Familiarity* 0.00 [0.00], and *Social Familiarity* 3.79 [0.00]. Since all the variables are statistically equivalent across the two samples, we can conclude that the exogenous assignment worked.[13]

Because we randomly select one team from each match to include, there is no guarantee that focal teams will win exactly half of the games in our sample. Indeed, as Table 1 shows, focal teams win 51% of all games in the sample. Specialist teams account for 10% of all teams,

Table 1. Descriptive Statistics and Pairwise Correlations for Key Variables

Variable		Mean	Std. dev.	Min	Max	1	2	3	4	5	6	7	8
1	*Victory*	0.52	0.50	0	1	1							
2	*Specialist Team*	0.08	0.26	0	1	0.01	1						
3	*Specialist Opponent*	0.08	0.27	0	1	−0.01	0.06	1					
4	*Generalist Team*	0.06	0.24	0	1	0.00	−0.07	0.00	1				
5	*Generalist Opponent*	0.06	0.24	0	1	0.00	0.00	−0.07	0.02	1			
6	*Social Familiarity*	28.58	84.99	0	3,966	0.03	0.02	0.01	0.05	0.00	1		
7	*Opponent Social Familiarity*	32.36	91.45	0	3.964	−0.05	0.00	0.02	0.00	0.05	0.04	1	
8	*Functional Familiarity*	0.53	0.13	0.02	0.8	−0.03	−0.04	−0.07	−0.17	−0.06	−0.23	0.05	1
9	*Opponent Functional Familiarity*	0.52	0.13	0.02	0.8	0.05	−0.07	−0.03	−0.06	−0.17	0.05	−0.24	0.40

Notes. In the regressions, *Social Familiarity* and *Opponent Social Familiarity* are standardized to be mean zero, and the standard deviation is equal to one. N = 6,444,502. Std. dev., Standard deviation.

whereas generalist teams account for 6% of all teams. *Functional Familiarity* has a mean of 0.55 and standard deviation of 0.12, a minimum of 0.1 and a maximum of 0.8 for both types of teams. The interpretation is that a typical dyad is about half functionally familiar with their teammates' hero portfolio.

The mean of *Social Familiarity* is 28.58, meaning that on average a player dyad within a team has played together about 28 times in the past. Although it may appear surprising that player dyads would have such frequent interaction, given that players are exogenously assigned to teams in our test sample, three factors tend to increase *Social Familiarity*. First, although players are exogenously assigned to teams in our test sample, social familiarity itself is not randomly determined. Indeed, our measure of *Social Familiarity* includes all matches played by every professional player, including matches where teams are endogenously determined (e.g., in tournaments). Thus, our social familiarity construct accurately captures how often players have played together. Second, exogenous assignment of players to teams takes place within player quality bins, so that stars play with other stars and not with novices and other nonstars. Professional players make up a tiny fraction of all DOTA2 players, but a significant fraction of star players, meaning they will end up playing together more often than any two randomly selected players from the full population of DOTA2 players. Third, players are exogenously assigned to teams of players who are on the platform at the same time. Although the modal DOTA2 player lives in East Asia, DOTA2 has a worldwide following (Merwin et al. 2018), meaning that time zones influence the "thickness" of the matching algorithm. Players entering the platform from a less popular time zone are far more likely to be assigned to a team together. As a result of endogenous team formation outside of the test sample, binning by ability, and time zone effects, certain players will tend to play much more often with one other; nevertheless, the assignment of a particular level of social familiarity to a team is exogenous.

Although social familiarity is exogenously assigned, there are clearly some extreme outliers in the data that should be considered carefully. In our robustness checks, we verify that the results are robust to Winsorizing the outliers. The results are also robust to trimming off outliers. We also account for the fact that the distribution is skewed to the right by standardizing the *Social Familiarity* to be mean zero with standard deviation one in the regressions.

One of the pairwise correlations is particularly noteworthy in Table 1. *Functional Familiarity* is negatively correlated with *Social Familiarity* at −0.23 for focal teams and −0.24 for opponent teams. Since we know players are exogenously assigned to teams within our test sample, those correlations are likely due to relationships in the data outside of the test sample, suggesting that endogenously formed teams with high social familiarity purposefully avoided forming teams with high functional overlap. Thus, using revealed preference to infer optimization, we might expect that the main effect of functional familiarity will be negative in our regressions.

Empirical Design

We conduct our main analyses at the team-match level. Our core test of the marginal effect of specialization and familiarity on performance is, for team j and match m:

$$\begin{aligned} Victory_{jm} = a &+ B_1 SpecialistTeam_{jm} + \boldsymbol{Familiarity_{jm}} \boldsymbol{B_{Fam}} \\ &+ \left(\boldsymbol{Familiarity_{jm}} \times SpecialistTeam_{jm}\right)\boldsymbol{B_{Own}} \\ &+ B_4 SpecialistOpponent_{jm} \\ &+ \left(\boldsymbol{Familiarity_{jm}} \times SpecialistOpponent_{jm}\right)\boldsymbol{B_{Opp}} \\ &+ \boldsymbol{X_c} \boldsymbol{B_c} + e_{jm}, \end{aligned} \tag{1}$$

where $\boldsymbol{c}$ indexes a vector of control variables. Boldface type indicates vector notation. The key coefficients of interest are $\boldsymbol{B_{Own}}$ and $\boldsymbol{B_{Opp}}$, which measure the complementarity between specialized teams and familiarity, where $\boldsymbol{Familiarity}$ includes both measures of familiarity: *Social Familiarity* and *Functional Familiarity*. In other regressions we modify expression (1) by exploring how the familiarity-specialization complementarity changes with fine-grained changes in team composition, using alternative definitions of specialization, and by including additional controls. The vector of controls in expression (1) includes the main and marginal effects of *Generalist Teams* and *Generalist Opponent* (i.e., teams where three or more members are generalists), and a full suite of month fixed effects. The excluded category is nonspecialist/nongeneralist teams. To rule out the possibility that certain team configurations influence the odds of winning, we also include a set of team configuration indicators that correspond to the number of carry heroes in each team. Finally, to rule out systematic temporal effects, we also included a full set of indicators that correspond to the year and calendar month in which the match takes place.

Since players are exogenously assigned to teams in the sample, none of the team-level independent variables in our regressions are choice variables at the team level. For example, even though individual specialization is endogenous to any individual, the measure *Specialist Team* is exogenous at the team level—the individual has no control over the team to which they are assigned. Given that the main effects of familiarity and team specialization are exogenous, the interaction terms, and opponent variables, are also exogenous (Athey and Stern 1998), and we can interpret the coefficients on the interaction terms as the causal effect of the complementarity between specialized teams and familiarity on team performance. However, it is important that we do not assume that statistical identification of a causal effect means that we have iron clad evidence for the mechanism in our theory. If, for example, *Specialist Team* is just a proxy for "highly skilled team," then the interpretation of the complementarity with familiarity would be different from our theory. The implication for this study is that although exogenous assignment of players to teams allows us to test directly whether there are complementarities between familiarity and specialization, we will need to consider the full gestalt of the evidence in order to conclude that extemporaneous coordination is the mechanism at work in the data. After presenting the econometric evidence, we discuss the interpretation of the results in more detail in the Discussion section.

Although our theory makes a simple, broad prediction—that familiarity and team specialization are complements in contexts where extemporaneous coordination is important—there are several testable implications of the theory. For example, besides predicting a positive marginal effect between team specialization and both social and functional familiarity for a focal team, the theory also suggests that we should see parallel, but opposite signed, effects for a focal team's opponent. As we demonstrate next, these predictions are also borne out in the data.

Main Results

For presentation purposes, all coefficient estimates and standard errors in the regression tables are multiplied by 100. To focus on the marginal effects of social familiarity, Table 2 represents a variation on expression (1), which excludes the marginal effects of functional familiarity. For a focal team, the main effect of specialization and social familiarity are positive (see Table 2). Column (1) shows that specialist teams are 1.75 percentage points more likely to win relative to nonspecialist/nongeneralist teams. Most importantly for our theory, the coefficient on the interaction between social familiarity and the specialist team dummy is 0.53 and is precisely estimated. The interpretation is that for a specialist team, a one

Table 2. Team Composition, Social Familiarity, and the Probability of Victory

	Number of specialists/generalists per team		
	3+	3+	4+
Variable	(1)	(2)	(3)
Dependent variable: *Victory*			
Specialist Team	1.75*	1.75*	2.38*
	(0.08)	(0.08)	(0.20)
Generalist Team	–0.35*	–0.35*	–0.95*
	(0.08)	(0.08)	(0.25)
Specialist Opponent	–2.30*	–2.28*	–3.09*
	(0.07)	(0.07)	(0.19)
Generalist Opponent	0.70*	0.64*	1.01*
	(0.08)	(0.08)	(0.25)
Social Familiarity	0.89*	0.90*	0.93*
	(0.02)	(0.02)	(0.02)
Opponent Team Social Familiarity	–1.79*	–1.80*	–1.81*
	(0.02)	(0.02)	(0.02)
Specialist Team × Social Familiarity	**0.53***	**0.56***	**0.75***
	(0.07)	**(0.07)**	**(0.16)**
Generalist Team × Social Familiarity	–0.15*	–0.16*	–0.30*
	(0.06)	(0.06)	(0.14)
Specialist Opponent × Opponent Social Familiarity		**–0.43***	**–0.19**
		(0.07)	**(0.17)**
Generalist Opponent × Opponent Social Familiarity		0.32*	0.25
		(0.05)	(0.13)
Controls	Yes	Yes	Yes

Notes. All regressions include *Functional Familiarity* and *Opponent Functional Familiarity* controls, as well as month fixed effects and team configuration indicators. For presentation purposes, all coefficients are multiplied by 100. Robust standard errors in parentheses. Specification (1): $Victory_{jm} = a + B_1 Specialist\ Team_{jm} + B_2 Social\ Familiarity_{jm} + B_3(Social\ Familiarity_{jm} * Specialist\ Team_{jm}) + \mathbf{X_c B_c} + e_{jm}$, where j indexes a team, m a match, and c controls. N = 6,444,502. Key results in bold.

*$p < 0.05$.

standard deviation increase in *Social Familiarity* is associated with a 0.53 percentage point increase in the probability of victory, relative to nonspecialist/nongeneralist teams with the same level of social familiarity. The marginal effect is larger still for specialist teams compared with generalist teams—0.68 percentage points (0.53 + 0.15), and precisely estimated (the t-statistic on the difference is 8.56). In other words, specialist teams benefit more from being more familiar with one another than nonspecialist teams, and particularly generalist teams, at the same level of social familiarity. Although the point estimates may not appear to be large, they are economically important. By design, each game is balanced such that the ex ante probability of winning is approximately 50%, so that even small changes in the probability of victory are meaningful. As expected, the coefficient estimates on the opponent team variables are in the opposite direction of the corresponding focal team variables, and of the same approximate economic magnitude (column (2)). Similar results are obtained running the same regression (as in column (2)) where specialist teams are teams with at least four specialists, though the point estimate on *Specialist Opponent* × *Opponent Social Familiarity* becomes indistinguishable from zero (column (3)).

To focus on the marginal effects of functional familiarity, Table 3 represents a variation on expression (1), which excludes the marginal effect of social familiarity. Column (1) of Table 3 shows that the main effect of functional familiarity is to decrease a focal team's chance of winning: a one standard deviation increase in *Functional Familiarity* decreases a focal team's chances of winning by 2.35 percentage points (column (1)). The negative coefficient is probably due to the fact that functional familiarity implies functional redundancy. Importantly for our theory, the interaction between functional familiarity and the specialist team dummy is (positive) 1.35 and statistically significant, meaning that a one standard deviation in *Functional Familiarity* increases a specialist team's probability of winning by 1.40 percentage points, compared with a nonspecialist/nongeneralist team at the same level of functional familiarity.

Table 3. Team composition, functional familiarity, and the probability of victory

	Number of specialists/generalists per team		
	3+	3+	4+
Variable	(1)	(2)	(3)
Dependent variable: *Victory*			
Specialist Team	1.95*	1.95*	2.63*
	(0.08)	(0.08)	(0.19)
Generalist Team	−0.34*	−0.34*	−0.73*
	(0.10)	(0.10)	(0.34)
Specialist Opponent	−2.29*	−2.46*	−3.14*
	(0.07)	(0.08)	(0.19)
Generalist Opponent	0.70*	0.89*	1.22*
	(0.08)	(0.10)	(0.33)
Functional Familiarity	−2.35*	−2.37*	−2.25*
	(0.02)	(0.02)	(0.02)
Opponent Functional Familiarity	2.83*	2.94*	2.81*
	(0.02)	(0.02)	(0.02)
Specialist Team* × *Functional Familiarity	**1.35***	**1.41***	**2.69***
	(0.09)	**(0.09)**	**(0.22)**
Generalist Team × *Functional Familiarity*	0.14	0.14	0.37
	(0.09)	(0.09)	(0.26)
Specialist Opponent* × *Opponent Functional Familiarity		**−1.85***	**−3.65***
		(0.09)	**(0.22)**
Generalist Opponent × *Opponent Functional Familiarity*		0.20*	0.12
		(0.09)	(0.26)
Controls	Yes	Yes	Yes

Notes. All regressions include *Social Familiarity* and *Opponent Social Familiarity* controls, as well as month fixed effects and team configuration indicators. For presentation purposes, all coefficients are multiplied by 100. Robust standard errors in parentheses. Specification (1): $Victory_{jm} = a + B_1 Specialist\ Team_{jm} + B_2 Functional\ Familiarity_{jm} + B_3(Functional\ Familiarity_{jm} * Specialist\ Team_{jm}) + \mathbf{X_c B_c} + e_{jm}$, where j indexes a team, m a match, and c controls. N = 6,444,502. Key results in bold.

*$p < 0.05$.

The marginal effect of functional familiarity is not precisely estimated for generalist teams. Still, compared with a generalist team, at the same level of functional familiarity, the marginal effect of functional familiarity on a specialist team is 1 percentage points, and precisely estimated (t-statistic on the difference of 9.51).

Column (2) of Table 3 includes the functional familiarity interaction terms for opponent teams, with similar results. A one standard deviation increase in a specialist opponent's functional familiarity decreases a focal team's probability of winning by a precisely estimated 1.85 percentage points, relative to a nonspecialist/nongeneralist opponent, at the same level of social familiarity. Compared to playing a generalist opponent, at the same level of functional familiarity, the marginal effect of functional familiarity on a specialist opponent team is 1.65 (1.85 – 0.20) percentage points.

Column (3) of Table 3 runs the same regression as in column (2), where specialized teams are redefined as teams with four or more specialists. The results are, again, similar in terms of signs and statistical significance, though the point estimates are somewhat larger. In sum, the results of Table 3 suggest that whereas increased functional familiarity is associated with a decreased chance of winning, specialist teams benefit from increased functional familiarity, relative to nonspecialist teams with the same level of functional familiarity.

Table 4 tabulates the full core specification represented by expression (1). It includes all the key marginal effects in a single specification. Column (1) shows the results when specialist teams are defined to have three or more specialists, and column (2) shows the results when specialist teams are defined to have four or more specialists. The results in both columns are consistent with Tables 2 and 3 and with one another. All of the key interactions are of the sign predicted, seven of the eight coefficients are precisely estimated, and the magnitudes are all reasonable and meaningful economically.

Table 4. Team composition and the probability of victory: full specification

	Number of specialists/ generalists per team	
	3+	4+
Variable	(1)	(2)
Dependent variable: *Victory*		
Specialist Team	1.92*	2.49*
	(0.08)	(0.2)
Generalist Team	–0.34*	–0.74*
	(0.10)	(0.34)
Specialist Opponent	–2.44*	–3.13*
	(0.08)	(0.19)
Generalist Opponent	0.90*	1.22*
	(0.10)	(0.33)
Social Familiarity	0.87*	0.92*
	(0.02)	(0.02)
Functional Familiarity	–2.37*	–2.25*
	(0.02)	(0.02)
Opponent Social Familiarity	–1.77*	–1.80*
	(0.02)	(0.02)
Opponent Functional Familiarity	2.94*	2.81*
	(0.02)	(0.02)
Specialist Team × Social Familiarity	**0.66***	**0.72***
	(0.07)	**(0.16)**
Specialist Team × Functional Familiarity	**1.48***	**2.67***
	(0.09)	**(0.22)**
Specialist Opponent × Opponent Social Familiarity	**–0.57***	**–0.17**
	(0.07)	**(0.17)**
Specialist Opponent × Opponent Functional Familiarity	**–1.90***	**–3.65***
	(0.09)	**(0.22)**
Controls	Yes	Yes

Notes. All regressions include *Generalist Team* and *Generalists Opponent* interactions with *Social Familiarity* and *Functional Familiarity*, month fixed effects and team configuration indicators. For presentation purposes, all coefficients are multiplied by 100. Robust standard errors in parentheses. $N = 6{,}444{,}502$. Key results in bold.

$^{*}p < 0.05$.

Taken together, Tables 2–4 suggest that familiarity and specialization exhibit complementarities. Figure 2 makes a related point representing a variation on expression (1), where the categorical variable *Specialist Team* is replaced with a set of categorical variables representing the number of specialists on a team. Adding specialists to the team (moving rightward in the figure)—making the team more specialized—tends to increase the marginal effect of familiarity. The pattern offers further evidence of the complementarity between specialization and familiarity.

Extensions and Robustness Checks

Individual Effects

Though our main focus is on team specialization and familiarity, we also found similar results at the individual level. Although individual characteristics are not exogenously assigned, specifications including player-specific fixed effects on the full sample of 64 million player-match observations yield results consistent with the team-level results (see the online companion for more details).

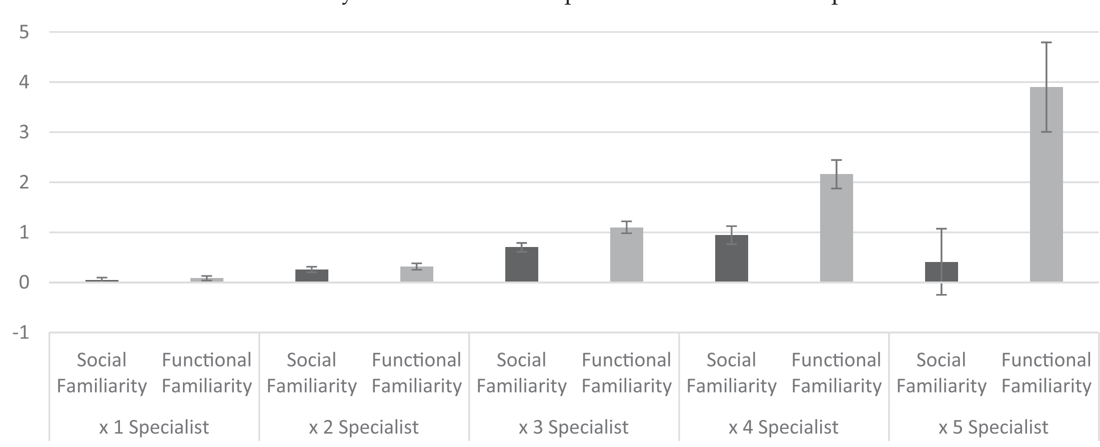
Alternative Measures of Specialization

We verify that the results are not sensitive to reasonable changes in the definition and thresholds for specialization, including: (i) requiring teams to have four or more specialists to be considered specialist teams, (ii) defining individual specialization more broadly to include all players above the median level of the individual specialization distribution, and (iii) allowing specialization to enter as a continuous variable instead of as a categorical variable. With respect to (i), specifications where teams with only four or more specialists are defined as specialized teams delivered similar results to our core specification, even returning larger point estimates for three of the four key interaction terms, though the interaction on *Specialist Opponent* × *Opponent Social Familiarity* becomes indistinguishable from zero when specialized teams are defined more strictly. With respect to (ii), setting the threshold for an individual to be defined as a specialist at the median of the distribution for individual specialization, instead of using the 75th percentile as the cut-off, also yielded similar results. However, as expected, using a broader definition of individual specialization does generally tend to attenuate the key coefficient estimates—three of the four point estimates were smaller, and *Specialist Opponent* × *Opponent Social Familiarity* became indistinguishable from zero. With respect to (iii), replacing the discrete measure of team specialization with a continuous measure, using the average of all teammates' *Individual Specialization*, led to only small changes in the point estimates, and all key interactions continued to be precisely estimated. The consistency of the results across a range of alternative measures of specialization (tabulated in the online companion), reinforces our key finding: specialization and familiarity are complements in team production.

Additional Controls

We perform a set of robustness checks that include additional controls. Although the specialization measure used in our empirical tests is computed for players with the exact same number of matches played, one still might reasonably wonder if controlling for overall team experience (i.e., the sum of the number of matches played by all four teammates) could influence the regression estimates. However, including controls for

Figure 2. Social and Functional Familiarity Interacted with a Spline of the Number of Specialists



Notes. The figure shows bar plots of the coefficients on the interactions for focal teams from regressions analogous to Table 4, column (1), with twice the standard error bars, but where the categorical variables *Specialist Team* and *Specialist Opponent* are replaced with a spline of categorical variables for teams with one, two, three, four, and five specialists, respectively. The vertical axis represents changes in the percentage change in the probability of victory for a team. Specification: $Victory_{jm} = a + \mathbf{Specialist\ Team_{jm}}\ \mathbf{B_{Spec,Own}} + \mathbf{Familiarity_{jm}}\ \mathbf{B_{Fam}}\ (\mathbf{Familiarity_{jm}} * \mathbf{Specialist\ Team_{jm}})\ \mathbf{B_{Own}} + \mathbf{Specialist\ Opponent_{jm}}\ \mathbf{B_{Spec,Opp}} + (\mathbf{Familiarity_{jm}} * \mathbf{Specialist\ Opponent_{jm}})\mathbf{B_{Opp}} + \mathbf{X_c}\ \mathbf{B_c} + \mathbf{e_{jm}}$, where j indexes a team, m a match, and c controls. Boldface indicates vector notation.

team experience does not meaningfully change the coefficient estimates on the four key interaction terms. The interpretation is that overall team experience is not an important omitted variable in our main analyses.

Since within-team familiarity is so important for focal and opponent teams, one might also wonder if familiarity with members of an opponent team matters, not through a coordination effect, but through an anticipation effect, whereby knowing opponents better allows one to anticipate their actions more accurately. There are two kinds of rival familiarity—having played with players previously that are now on the opposing team, and having played against players previously that are now on a rival team. However, including either or both forms of rival familiarity did not change the main results, suggesting that rival familiarity is not an important omitted variable in our main analyses, and, moreover, that the complementarity between specialization and familiarity flows through a coordination mechanism, and not simply an anticipation mechanism.

In approximately 7% of the matches in our test sample, at least one player abandons the match before it is complete. However, controlling for those matches directly in the regression yield similar to those in our core specification, suggesting that abandoned games do not bias the results. We also verify that including a full set of hero fixed effects to control for hero-specific sources of heterogeneity does not materially change results.

Alternative Specifications and Samples

We verify that the main results are not biased from nonrandom assignment of some team members in a small subset of ranked matches. During our observation period, DOTA2 experimented with a party system that occasionally allowed groups of two or three players to join matches together. Fortunately for our research design, the party system was only deployed by the game publisher in a limited way. Unfortunately, DOTA2 only explicitly flagged party matches beginning in April 2016. To verify that party matches do not bias the results, we restrict the sample to the 195,374 matches occurring after April 2016 (to December 2016 when we end our data collection). Of these matches, 169,918 explicitly do not feature parties, (i.e., all 10 players are exogenously assigned). Running our main specification on the 169,918 non-party matches led to point estimates on the key coefficients that are larger, as are the standard errors, but all remain statistically significant and of the same sign. The results suggest that limited party play did not meaningfully bias the results in the full sample.

We also verify that our main specification has meaningful predictive power. We begin by randomly selecting a subset of matches as our training sample, keeping the rest as a testing sample. We then predict the outcome of the matches in the testing sample, using estimates drawn from regressions performed on the training sample. We repeat this exercise 100 times. The results are striking. Based on the estimates from our training sample, our main specification is significantly (three to five percentage points) better than the baseline prediction one would expect from the DOTA2 algorithm (please see the online companion for more details). The predictive power of our specification gives us added confidence that our specification captures real economic effects.

Discussion

In our empirical tests, research design facilitates causal inference. The results show specialization and familiarity are complements. But what about the mechanism? Is the complementarity due to extemporaneous coordination, as in our theory, or is something else at work? Unfortunately, we cannot observe extemporaneous coordination directly in the data, and so we must use theory, contextual evidence, and different cuts of the data to eliminate alternative hypotheses. The strongest evidence for the extemporaneous coordination mechanism comes from the institutional details of the game. As in many traditional sports settings, DOTA2 requires extensive coordination on the fly. Indeed, it would not be overstating the case to say extemporaneous coordination is crucial for success in DOTA2. However, we know ex ante coordination over the division of labor—particularly role selection—matters too.

The effect of selection into roles (i.e., hero selection) represents the strongest rival mechanism for explaining complementarity between specialization and familiarity. And, indeed, we cannot rule out the possibility that the marginal effect of social familiarity with specialist teams represents a combination of ex ante and extemporaneous coordination. However, theory and institutional details suggest that the marginal effect of functional familiarity on specialist teams, relative to generalist teams, represents a clean test in favor of the extemporaneous coordination mechanism. To see why, first recall that the cost of functional familiarity is redundancy, a cost that is represented by the negative coefficient estimate on the main effect of functional familiarity in the regressions (e.g., column (1) of Table 4). Second, from theory and observation, we know that on the margin, redundancy costs are higher for specialist teams than for generalist teams; generalist teams have several players that can play many roles, and so can more easily adjust to being assigned a team with redundancies than specialist teams. Thus, the marginal effect of functional familiarity and specialization can only be positive, relative to generalist teams, if the extemporaneous coordination benefits of familiarity outweigh the costs of redundancy. The tests show exactly that the marginal effect of functional familiarity on team composition is statistically greater for specialist teams than for generalist teams (e.g., from Table 3, column (2) the t-statistic on the difference between the mean of *Specialist Team* × *Functional Familiarity* and the mean of *Generalist Team* × *Functional Familiarity* is 9.98). In other words, even though generalist teams are more effective at ex ante adaptation, in-game coordination effects associated with the marginal effect of functional familiarity and specialist team swamp the ex ante effects, suggesting that improved extemporaneous coordination drives the difference.

Our main results are about complementarities—interactions between team composition and familiarity. However, main effects are also important to understand. Given that specialist effects are strongly positive in the team-level regressions, a natural question arises: why isn't everyone a specialist in DOTA2? Taking the question one step further, one should also ask: does the strong specialist effect undercut the theory or identification strategy offered in this paper? Some players might become generalists because the heroes they specialized in became less valuable due to changes in game rules over time. Others might change their role in the hopes of filling a niche on a team they would like to play with in a tournament, whereas still others, out of curiosity, or perhaps boredom, might explore different roles more than others. Although we have no reason to think there is systematic sorting into specialists or generalists by ability—and indeed the within-person correlation between increased individual specialization and probability of victory is indistinguishable from zero (see the online companion for more details)—it is possible that an unobservable process could lead to a correlation between person-specific quality and team composition. However, even if nonspecialists are systematically lesser players, the main effect of team composition (e.g., *Specialist Team*) acts as a fixed effect that filters out the average effect of sorting, allowing for a clean interpretation of the marginal effects of specialist team with familiarity.

Although we offer large sample empirical evidence of complementarities between specialization and familiarity in a context where extemporaneous coordination is of great importance, this paper does have two meaningful limitations. A sceptic might reasonably say the results only represent six million experiments in the same context. To better understand how well the theory generalizes to other contexts, it should be tested in other settings. Also, since we do not observe extemporaneous coordination directly in the data, but rather infer coordination effects from theory, context-specific characteristics of the setting, and patterns in the data, future research on extemporaneous coordination in teams would benefit from additional microanalytic analysis of the inner workings of teams.

Conclusion

This paper studies team composition complementarities in contexts where extemporaneous coordination is important, advancing the idea that specialist teams will be relatively more effective when they are more familiar with their teammates. To test this proposition, we analyze millions of e-sports matches where teammates are exogenously assigned. We find that team specialization and familiarity are complements—the returns to specialization increase with familiarity, relative to nonspecialist teams with the same level of familiarity.

The paper offers new insights into organizational design. Increased worker mobility and the diffusion of project work is altering how organizations manage human capital and how individuals manage their careers. A number of studies report positive returns to specialization for individuals, suggesting there are strong incentives for individual workers to specialize. Our results show that familiarity between team members is an important catalyst for increasing the returns to teams of specialists in organizational contexts characterized by nondecomposable tasks and unpredictable interdependencies among team members.

Previous studies have highlighted the importance of coordinators in teams (e.g., Briscoe and Rogan 2016), and have questioned the ability of specialists to become efficient coordinators (e.g., Reyt et al. 2016). Our study builds on and extends this stream of research by showing that familiarity substitutes for breadth of task knowledge, allowing teams of specialists to become relatively more effective when they and their teammates have interacted more extensively, or have performed similar roles in the past.

Insight into the effects of specialization and teamwork is especially important today as the employment relationship is changing in many industries (Bidwell et al. 2013). Processes traditionally associated with middle managers—such as selection, team formation, and the allocation of personnel to specific tasks—are shifting (Mitchell and Brynjolfsson 2017), leading to increased extemporaneous coordination among teammates. The implications of such transformations for individual career trajectories as well as team assembly and outcomes are not fully explained by current theory.

Moreover, understanding how to design teams that can effectively coordinate when facing unpredictable interdependencies is a subject of growing importance. There are many contexts in which team composition and team familiarity vary meaningfully, including teams of management consultants, medical teams, film crews, sports teams, and entrepreneurial teams. These contexts are often characterized by considerable flux in team composition, as well as opportunities for repeated interaction between teammates, with individual discretion contributing to the successful completion of interdependent tasks. Our study offers new insights into how to improve the performance of such teams, which contributes to the scholarship and practice of optimal team design.

Acknowledgments
The authors would like to thank Colin Fisher, Brent Goldfarb, Martin Kilduff, Blaine Landis, John Mathieu, Damon Phillips, Katherine Phillips, Alessandro Piazza, Jonah Rawley, Ray Reagans, Bart Vanneste, Joel Waldfogel, Dan Wang, and seminar participants at University College London, University of Connecticut, University of Minnesota, Purdue University, Worcester Polytechnic Institute, Johns Hopkins University, IESEG School of Management, Fordham University, Manhattan College, the Academy of Management Conference, the BYU Winter Strategy Conference, the Institutions & Innovation Conference and the Strategic Management Conference for comments, and suggestions. The data in this paper were assembled with the excellent assistance of Jacqueline Araya Rebolledo, Jimmy Baodong Pan, James Kainesley and Daniel Mechanic. The authors are especially grateful to UCL School of Management and Worcester Polytechnic Institute for technical support. All errors remain their own. All authors contributed equally to this manuscript and are listed in alphabetical order.

Endnotes
[1] Task interdependencies exist where the efficiency of one task depends on another one (Natividad and Rawley 2016). Similarly, team member interdependencies exist where the efficiency of one team member depends on another. Teams experience interdependencies in many settings where sequences and allocations of tasks cannot be articulated with enough specificity and precision ex ante to allow for individuals to independently complete their work efficiently.

[2] Decomposability refers to the distribution pattern of interdependencies in task systems (Rivkin and Siggelkow 2007). "A task system is highly decomposable when its tasks can be divided into discrete subsystems, with dense interdependencies within subsystems and sparse interdependencies between them. For a given level of complexity, the more decomposable the task system, the easier it is to modularize and distribute coordination responsibilities across divisions" (Zhou 2013, p. 339). By contrast, the benefits of modularity are lower when task systems are nondecomposable.

[3] Scholars have reported positive effects of social familiarity on team processes and outcomes in a number of settings, such as surgery teams, software teams, sports teams, movie production teams, and SWAT teams. However, prior studies do not address whether specialization and familiarity are complements or substitutes.

[4] DOTA2 is a multiplayer online game developed and published by Valve Corporation. Each match pits two teams against each other in a battle to achieve a set of objectives. The game requires extensive coordination between the players, as well as considerable individual skill. The game is frequently referred to as one of the most complex computer games, and has been used for artificial intelligence challenges due to its sheer complexity. Each match lasts on average around 30 to 45 minutes though the longer matches can take more than an hour. See http://dota2.com for more information on the game mechanics.

[5] Players are assigned from a queue of players who join the gaming platform at approximately the same time. Games are stratified within bands of player ability. In some matches, groups of players may enter the gaming platform together. See Dota Team (2013) for more information on the DOTA2 match-making system.

[6] We divide heroes into two broad groups based on the game publisher's characterization. For more details on heroes' roles, see, for example, https://www.it-support.com.au/beginners-guide-to-dota-2-team-roles-part-1/2014/10.

[7] For example, a carry player is similar to a striker in soccer, whereas a support player is more like a defender. Analogous roles in a traditional firm might be front office versus back office positions.

[8] Although the exact ranking calculation Matchmaking Rating (MMR) is not published, it is a rating engine that assigns a score to players reflecting their approximate playing ability. Each player is assigned a rating that is dynamically updated based on match results (i.e., winning or losing). Within a band of MMR, players are considered to be approximately equal in playing ability—ranked matches are organized within these bands—and players are randomly assigned to teams (Dota Team 2013).

[9] Team play, where players pre-arrange to play together, is a relatively unusual format, representing about 2% of all ranked matches. Since teams are not exogenously assigned in team play, we exclude these matches from the analyses (though we include them for the purposes of calculating specialization and familiarity). In party play, groups of two or three may pre-arrange to be assigned to games together. Although DOTA2 has experimented with allowing limited amounts of party play over time, they unfortunately did not track party play games separately until April 2016. Thus, our baseline analyses do include some games where teams are not fully exogenously assigned, though we show that the results are robust to tests that explicitly exclude party play games in using the post-April 2016 subsample.

[10] DOTA2 provides a WebAPI where developers can retrieve complete match history and details. Players can hide their identity from the general public before the start of any match by playing anonymously. However, since serious players monetize their fan base by allowing fans to follow their matches, they are unlikely to play a significant number of games anonymously.

[11] Our sampling strategy represents a standard approach to estimating network measures from large networks when calculating measures on the full population would require extraordinary levels of computational power. The sampling strategy results in complete networks for all focal players (egos), but incomplete networks for nonfocal players (alters). However, since unobserved alter nodes are likely to be missing at random, and we have a large network, our sampling strategy should not meaningfully bias our estimates (Wang et al. 2012, Smith and Moody 2013, Smith et al. 2017).

[12] The results are robust to using the full sample, and to excluding all games with even a single anonymous player. We exclude unranked matches, because they do not affect players' ratings, and therefore may influence player effort. Finally, we also exclude from the test sample all matches where players are not exogenously assigned to teams (i.e., professional and team play matches where teams are endogenously determined).

[13] We also verify that the empirical distribution of the key explanatory variables is statistically indistinguishable from a simulated distribution where teams were randomly assigned (see the online companion for more details).

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

Athey S, Stern S (1998) An empirical framework for testing theories about complementarity in organizational design. Working paper, Stanford University, Stanford, CA.

Barley SR, Bechky BA, Milliken FJ (2017) The changing nature of work: Careers, identities, and work lives in the 21st century. *Acad. Management Discoveries* 3(2):111–115.

Bechky BA (2006) Gaffers, gofers, and grips: Role-based coordination in temporary organizations. *Organ. Sci.* 17(1):3–21.

Bechky BA, Chung DE (2018) Latitude or latent control? How occupational embeddedness and control shape emergent coordination. *Admin. Sci. Quart.* 63(3):607–636.

Bechky BA, Okhuysen GA (2011) Expecting the unexpected? How SWAT officers and film crews handle surprises. *Acad. Management J.* 54(2):239–261.

Becker GS (1962) Investment in human capital: A theoretical analysis. *J. Political Econom.* 70(5, Part 2):9–49.

Ben-Menahem SM, von Krogh G, Erden Z, Schneider A (2016) Coordinating knowledge creation in multidisciplinary teams: Evidence from early-stage drug discovery. *Acad. Management J.* 59(4):1308–1338.

Berman SL, Down J, Hill CWL (2002) Tacit knowledge as a source of competitive advantage in the national basketball association. *Acad. Management J.* 45(1):13–31.

Bidwell M, Briscoe F, Fernandez-Mateo I, Sterling A (2013) The employment relationship and inequality: How and why changes in employment practices are reshaping rewards in organizations. *Acad. Management Ann.* 7(1):61–121.

Briscoe F, Rogan M (2016) Coordinating complex work: Knowledge networks, partner departures, and client relationship performance in a law firm. *Management Sci.* 62(8):2392–2411.

Cattani G, Ferriani S, Frederiksen L, Täube F (2011) Project-based organizing and strategic management: A long-term research agenda on temporary organizational forms. *Adv. Strategic Management* 28:xv–xxxix.

Choudhury P, Haas MR (2018) Scope vs. speed: Team diversity, leader experience, and patenting outcomes for firms. *Strategic Management J.* 39(4):977–1002.

Clement J (2019) Sensing the next game changer: Organizational architecture, cognition and adaptation to environmental change. Working paper, Stanford University, Stanford, CA.

Deming DJ (2017) The growing importance of social skills in the labor market. *Quart. J. Econom.* 132(4):1593–1640.

Dota Team (2013) Matchmaking. *DOTA2* (December 6), http://blog.dota2.com/2013/12/matchmaking/.

Ferguson JP, Hasan S (2013) Specialization and career dynamics: Evidence from the Indian Administrative Service. *Admin. Sci. Quart.* 58(2):233–256.

Fonti F, Maoret M (2016) The direct and indirect effects of core and peripheral social capital on organizational performance. *Strategic Management J.* 37(8):1765–1786.

Gardner HK, Gino F, Staats BR (2012) Dynamically integrating knowledge in teams: Transforming resources into performance. *Acad. Management J.* 55(4):998–1022.

Gibbons R, Waldman M (2004) Task-specific human capital. *Amer. Econom. Rev.* 94(2):203–207.

Huckman RS, Staats BR, Upton DM (2009) Team familiarity, role experience, and performance: Evidence from Indian software services. *Management Sci.* 55(1):85–100.

Jarvenpaa SL, Leidner DE (1999) Communication and trust in global virtual teams. *Organ. Sci.* 10(6):791–815.

Kuhn KM, Maleki A (2017) Micro-entrepreneurs, dependent contractors, and instaserfs: Understanding online labor platform workforces. *Acad. Management Perspect.* 31(3):183–200.

Lazear EP (2004) Balanced skills and entrepreneurship. *Amer. Econom. Rev.* 94(2):208–211.

Lewis K, Lange D, Gillis L (2005) Transactive memory systems, learning, and learning transfer. *Organ. Sci.* 16(6):581–598.

Lewis K, Belliveau M, Herndon B, Keller J (2007) Group cognition, membership change, and performance: Investigating the benefits and detriments of collective knowledge. *Organ. Behav. Human Decision Processes* 103(2):159–178.

Luciano MM, Bartels AL, D'Innocenzo L, Maynard MT, Mathieu JE (2018) Shared team experiences and team effectiveness: Unpacking the contingent effects of entrained rhythms and task characteristics. *Acad. Management J.* 61(4):1403–1430.

Majchrzak A, Jarvenpaa SL, Hollingshead AB (2007) Coordinating expertise among emergent groups responding to disasters. *Organ. Sci.* 18(1):147–161.

Merwin CD, Sugiyama M, Mubayi P, Hari T, Terry HP, Duval A, Bellini H, et al. (2018) eSports from wild west to mainstream. Report, Goldman Sachs, New York.

Mitchell T, Brynjolfsson E (2017) Track how technology is transforming work. *Nature* 544(7650):290–292.

Narayanan S, Balasubramanian S, Swaminathan JM (2009) A matter of balance: Specialization, task variety, and individual learning in a software maintenance environment. *Management Sci.* 55(11): 1861–1876.

Natividad G, Rawley E (2016) Interdependence and performance: A natural experiment in firm scope. *Strategy Sci.* 1(1):12–31.

Okhuysen GA, Bechky BA (2009) 10 Coordination in organizations: An integrative perspective. *Acad. Management Ann.* 3(1):463–502.

Postrel S (2002) Islands of shared knowledge: Specialization and mutual understanding in problem-solving teams. *Organ. Sci.* 13(3):303–320.

Puranam P, Raveendran M, Knudsen T (2012) Organization design: The epistemic interdependence perspective. *Acad. Management Rev.* 37(3):419–440.

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

Reagans R, Miron-Spektor E, Argote L (2016) Knowledge utilization, coordination, and team performance. *Organ. Sci.* 27(5):1108–1124.

Ren Y, Argote L (2011) Transactive memory systems 1985–2010: An integrative framework of key dimensions, antecedents, and consequences. *Acad. Management Ann.* 5(1):189–229.

Reyt JN, Rubineau B, Wiesenfeld BM (2016) The effect of resumes' level of concreteness on candidate evaluation. *Acad. Management Proc.* 2016(1):16694.

Rico R, Sánchez-Manzanares M, Gil F, Gibson C (2008) Team implicit coordination processes: A team knowledge–based approach. *Acad. Management Rev.* 33(1):163–184.

Rivkin JW, Siggelkow N (2007) Patterned interactions in complex systems: Implications for exploration. *Management Sci.* 53(7): 1068–1085.

Schakel JK, van Fenema PC, Faraj S (2016) Shots fired! Switching between practices in police work. *Organ. Sci.* 27(2):391–410.

Schilling MA, Vidal P, Ployhart RE, Marangoni A (2003) Learning by doing something else: Variation, relatedness, and the learning curve. *Management Sci.* 49(1):39–56.

Skilton PF, Dooley KJ (2010) The effects of repeat collaboration on creative abrasion. *Acad. Management Rev.* 35(1):118–134.

Smith A (1776) Campbell R, Sinner A, eds. *An Inquiry into the Nature and Causes of the Wealth of Nations* (Oxford University Press, New York).

Smith EB, Hou Y (2015) Redundant heterogeneity and group performance. *Organ. Sci.* 26(1):37–51.

Smith JA, Moody J (2013) Structural effects of network sampling coverage I: Nodes missing at random. *Soc. Networks* 35(4):652–668.

Smith JA, Moody J, Morgan JH (2017) Network sampling coverage II: The effect of non-random missing data on network measurement. *Soc. Networks* 48:78–99.

Sohn MW (2001) Distance and cosine measures of niche overlap. *Soc. Networks* 23(2):141–165.

Srikanth K, Harvey S, Peterson R (2016) A Dynamic Perspective on Diverse Teams: Moving from the Dual-Process Model to a

Dynamic Coordination-based Model of Diverse Team Performance. *Acad. Management Ann.* 10(1):453–493.

Staats BR, Gino F (2012) Specialization and variety in repetitive tasks: Evidence from a Japanese bank. *Management Sci.* 58(6):1141–1159.

Stan M, Puranam P (2017) Organizational adaptation to interdependence shifts: The role of integrator structures. *Strategic Management J.* 38(5):1041–1061.

Stuart HC (2017) Structural disruption, relational experimentation, and performance in professional hockey teams: A network perspective on member change. *Organ. Sci.* 28(2):283–300.

Taylor F (1911) *The Principles of Scientific Management* (Harper Brothers, New York).

Teodoridis F (2018) Understanding team knowledge production: The interrelated roles of technology and expertise. *Management Sci.* 64(8):3625–3648.

Teodoridis F, Bikard M, Vakili K (2019) Creativity at the knowledge frontier: The impact of specialization in fast- and slow-paced domains. *Admin. Sci. Q.* 64(4):894–927.

Valentine MA, Edmondson AC (2015) Team scaffolds: How mesolevel structures enable role-based coordination in temporary groups. *Organ. Sci.* 26(2):405–422.

Waguespack DM, Dunford ET, Birnir JK (2018) Cultural imprinting, institutions, and the organization of new firms. *Strategy Sci.* 3(2):426–438.

Wang DJ, Shi X, McFarland DA, Leskovec J (2012) Measurement error in network data: A re-classification. *Soc. Networks* 34(4):396–409.

Zhou YM (2013) Designing for complexity: Using divisions and hierarchy to manage complex tasks. *Organ. Sci.* 24(2):339–355.

Zuckerman EW (2017) The categorical imperative revisited: Implications of categorization as a theoretical tool. Durand R, Granqvist N, Tyllström A, eds. *From Categories to Categorization: Studies in Sociology, Organizations and Strategy at the Crossroads*, vol. 51 (Emerald Publishing Limited, Bingley, UK), 31–68.

Zuckerman EW, Kim T, Ukanwa K, von Rittmann J (2003) Robust identities or nonentities? Typecasting in the feature-film labor market. *Amer. J. Sociol.* 108(5):1018–1073.

Kenny Ching is an assistant professor of strategy and entrepreneurship and a member of the graduate faculty at Worcester Polytechnic Institute. His research interests include team performance, entrepreneurial strategy, and sports analytics, with particular emphasis on applying big data and machine learning technologies.

Enrico Forti is an assistant professor of strategy at Manhattan College's O'Malley School of Business and a lecturer at the UCL School of Management, University College London. He is also a Chazen Adjunct Research Scholar at Columbia Business School, Columbia University. His research draws on organizational theory to study the ways that strategy and organizational structure interact to influence performance. He is particularly interested in the relationship between firm scope and performance.

Evan Rawley is an associate professor of strategic management and entrepreneurship at the University of Minnesota's Carlson School of Management. His research interests include corporate strategy, entrepreneurship, and organizational economics.