HARVARD | BUSINESS | SCHOOL

Performance Tradeoffs in Team Knowledge Sourcing

Bradley R. Staats
Melissa A. Valentine
Amy C. Edmondson

Working Paper

11-031

October 14, 2011

Performance Tradeoffs in Team Knowledge Sourcing

Bradley R. Staats
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-3490
Tel: 919.962.7343
bstaats@unc.edu

Melissa A. Valentine
Harvard Business School
Boston, MA 02163
Tel: 617.852.8644
mvalentine@hbs.edu

Amy C. Edmondson
Harvard Business School
Boston, MA 02163
Tel: 617.495.6732
aedmondson@hbs.edu

October 14, 2011

Abstract

This research examines how teams organize *knowledge sourcing* (obtaining access to others' knowledge or expertise) and investigates the performance trade-offs involved in two approaches to knowledge sourcing in teams. One approach a team can take is to specialize, such that a small number of members source knowledge on behalf of the team. This *specialized* knowledge-sourcing approach lowers search costs. The other approach has most or all team members engaging in knowledge sourcing. This *broad* approach means that more team members interact directly with the knowledge source, and thus may understand the knowledge better. These options present a sourcing paradox: teams cannot reap the advantages of specialized sourcing *and* the advantages of broad sourcing. They face performance tradeoffs. Further under some conditions performance tradeoffs will be more pronounced. Specifically, specialized knowledge sourcing depends on within team knowledge sharing, and so conditions that hinder knowledge sharing in a team are likely to reduce the effectiveness of the specialized approach. Using archival data from several hundred software development projects in an Indian software services firm, we find support for most of our hypotheses. Our findings offer insight for theory and practice into how team organization, organizational knowledge resources, and within-team knowledge sharing can aid team performance.

Key Words: Knowledge Management, Knowledge Sharing, Knowledge Sourcing, Learning, Performance Tradeoffs, Teams

1. Introduction

To cope with the demands of dynamic global markets, many organizations use project teams to design, develop, or deliver products and services. Project teams have the potential to recognize and adapt to changing situations as they unfold (Edmondson and Nembhard 2009; Huckman, Staats and Upton 2009; Hackman and Katz 2010). Teams can manage complex interdependencies and promote innovation and creativity (Ancona and Bresman 2007; Larson 2010). Yet, despite the known potential benefits, teams do not always perform well. In particular, the extent to which a team has the knowledge it needs to do its work has long been considered crucial to effective team performance (Hackman 1986). And access to accurate, timely knowledge is especially critical in today's dynamic environments (Cummings 2004; Haas 2006). Often the relevant knowledge may be held somewhere else in the organization than within the team, meaning the team must engage in *knowledge sourcing* – defined as accessing the knowledge and expertise of other organizational members – to acquire needed knowledge (Gray and Meister 2004). Understanding how teams acquire the knowledge that they need to do their work, and investigating whether different approaches to knowledge sourcing matter for team performance, are therefore important topics for research.

Prior research on knowledge sourcing has focused on whether and why individuals access knowledge and expertise from other organizational members (Gray and Meister 2004). However, the need to access organizational knowledge is a not limited to individuals; teams also need to access knowledge relevant to their work. *Team knowledge sourcing* is more complex than individual knowledge sourcing. For one, teams confront a choice about how many members of the team should source knowledge. A team can devote a small number of members to the task of knowledge sourcing, which we refer to as *specialized* knowledge sourcing. Alternatively, most or all team members may engage in knowledge sourcing, which we refer to as *broad* knowledge sourcing.

Both approaches to organizing team knowledge sourcing offer certain advantages. With specialized knowledge sourcing, one or a few team members gain expertise in how to locate relevant knowledge. A specialized approach prevents other team members from accruing search costs and having to translate external knowledge for use by the team, thereby buffering these team members from information processing demands, analogous to arguments advanced by Tushman (Tushman 1979; Tushman and Katz 1980) about gatekeepers in product development projects. In contrast, broad knowledge sourcing means that most or all team members engage in the process of searching for, accessing, transferring, and applying external knowledge. This approach to knowledge sourcing is more likely to result in improved awareness and understanding throughout a team of the knowledge that exists within the organization. A broad approach allows more team members to reflect on the sourced knowledge and to use it to adjust their understanding of their focal problem. Team members can then create new knowledge that integrates the sourced knowledge with their new understanding of the problem.

Teams will likely vary in terms of whether they use a specialized or broad approach, implicitly choosing between the efficiency benefits of a specialized sourcing approach and the quality benefits of a broad sourcing approach. The impossibility of pursuing both specialized and broad approaches – thus gaining both efficiency and quality – constitutes a knowledge sourcing paradox for teams. Further, we consider conditions that exacerbate the sourcing paradox. Specialized sourcing, as compared to broad

sourcing, depends on effective knowledge sharing within a team; the sourcing experts must transfer the knowledge they sourced to other team members. Therefore, under conditions that hinder knowledge sharing, specialized sourcing will be less effective. We examine the relationship between specialized sourcing and performance under three conditions that challenge within-team knowledge sharing: geographic dispersion, lack of team familiarity (i.e., members' prior shared work experience), and the presence of subgroups (Gibson and Vermeulen 2003; Huckman et al. 2009; O'Leary and Mortensen 2010).

Knowledge can be obtained from varied sources, including from an organizational knowledge repository (Stein and Zwass 1995; Gray and Meister 2004). A knowledge repository is an information system with a user interface, such as a corporate intranet, wherein existing organizational knowledge can be catalogued, searched, and accessed. Individuals in teams can access previously created and codified knowledge to solve their in-progress problems (Davenport and Prusak 1998; Hansen, Nohria and Tierney 1999; Alavi and Leidner 2001). We obtained archival data on team patterns of knowledge sourcing from a knowledge repository at Wipro Technologies, a global, outsourced provider of software services. Sourcing in the repository was tracked for over 9,000 individuals in more than 300 software project development teams on a per-click basis. We linked these empirical data on knowledge sourcing with other Wipro databases – including data on archival team performance outcomes and project characteristics – for software development projects completed during 2008 and 2009.

In this study we make several contributions to the literature on teams and knowledge management. First, we adapt the construct of knowledge sourcing (Gray and Meister 2004) to the team level of analysis. Second, we examine the performance advantages of specialized versus broad sourcing within teams. In so doing, we show that there is a trade-off inherent in knowledge sourcing approaches. Because the two sourcing strategies are mutually exclusive, a single team cannot gain the performance advantages offered by both. Our results also contribute to the literatures on geographically-dispersed teams, team familiarity, and subgroups effects on team knowledge sharing.

2. Knowledge Sourcing in Teams

Team sourcing from an organizational knowledge repository can improve team performance in several ways.[1] First, teams can replicate existing knowledge by applying it in new contexts or to new problems (March 1991; Gray and Meister 2004). Sourced knowledge also provides an alternative lens through which prior knowledge and existing problems can be viewed (Daft and Weick 1984), so that teams can interpret and adapt the knowledge to generate entirely new solutions to solve existing problems (Henderson and Clark 1990; Fleming and Sorenson 2004; Majchrzak, Cooper and Neece 2004). Empirical studies support the idea that knowledge sourcing from a repository can involve replication, interpretation or adaptation of knowledge, and show that team knowledge sourcing is generally associated with better team performance (e.g., Haas and Hansen 2005; Haas and Hansen 2007). These prior studies of knowledge sourcing and team performance have considered the mean level of sourcing within a team. We build on this theory by hypothesizing that even holding the level of knowledge sourcing constant, how sourcing is organized will affect performance. Specifically, we examine whether and how specialized and broad team knowledge sourcing affects team performance.

2.1 Organizing Team Knowledge Sourcing: Specialized versus Broad Sourcing

The process of sourcing and then integrating sourced knowledge into the team deliverable looks significantly different in a specialized versus broad sourcing pattern. The main difference is in where the majority of the information processing of externally sourced knowledge occurs. Each sourcing pattern involves search, transfer and application of knowledge, but the differences in where the information processing occurs have implications for how each sourcing pattern is likely to affect team performance.

The search process will be more efficient in a specialized sourcing pattern. Knowledge search is an uncertain process with an objective of rapidly and accurately identifying a good solution; experience thus improves search effectiveness (Simon 1962; Nickerson and Zenger 2004). Sourcing experts in a

---

[1] A line of research has examined predictors of the creation of knowledge for storage in an organizational repository (Alavi and Leidner 2001; Cabrera and Cabrera 2002; Bock, Sabherwal and Qian 2008). Researchers have found that

team will likely develop fluency in the technical language of the system after repeated experience structuring queries. Learning how to structure queries effectively increases search efficiency (Hansen and Haas 2001; Gray and Durcikova 2005). With specialists, fewer people accrue search costs, and those who search will be those whose costs are lowest.

Specialized knowledge sourcing also may improve the efficiency of the knowledge transfer process. Szulanski (1996; 2000) showed that even a motivated receiver of new knowledge may struggle to understand it, without sufficient expertise (Cohen and Levinthal 1990; Boh 2008). Thus, having learned the technical language of a repository, a sourcing expert may be better at understanding and transferring knowledge from the repository to the team by translating or clarifying the knowledge.

Finally, a specialized sourcing pattern may improve the efficiency of the application process. Knowledge application can occur through replication, wherein the newly gained knowledge is exploited, or simply used as-is; or through adaptation, wherein the newly gained knowledge is changed or combined with other knowledge to generate a new solution (March 1991; Gray and Meister 2004). Tushman and Katz (1983) argue that teams that use a gatekeeper approach will be more likely to replicate, rather than adapt the knowledge that they acquire (see also, Tushman 1977; Allen, Tushman and Lee 1979; Katz and Tushman 1979; Allen, Lee and Tushman 1980). A gate keeper buffers team members from having to process external information. As a result, team members are more likely to simply replicate the sourced knowledge rather than making sense of it in the old context, understanding its relevance for the new context, and transforming it into a new solution.

Each part of the knowledge search, transfer, and application process are made more efficient (e.g., are more likely to be delivered on budget) by a specialized sourcing approach. As a result, we hypothesize:

HYPOTHESIS 1: *Specialized team knowledge sourcing is positively associated with project efficiency performance.*

Broad knowledge sourcing, on the other hand, involves many or all team members engaging in information processing of externally sourced knowledge. This dynamic will also influence the search, transfer, and application of the knowledge. Moreover, broad sourcing depends less on within team knowledge transfer. These differences will influence how a broad sourcing pattern influences performance.

In a team with a broad sourcing approach, the search is conducted by the team member who needs the knowledge, rather than by the sourcing expert who is one-step removed from the problem context. Individuals who directly source knowledge may encounter new ideas or problem-solving strategies, which they can combine with their existing knowledge, possibly helping them solve their current problems. By way of analogy, consider conducting a literature review for a joint academic paper. Although a skilled assistant might efficiently conduct the review on behalf of all of the authors, the quality of insight is likely to be enhanced by each author doing a search, allowing new ideas to be identified and disparate threads to come together in the process. This might hinder efficiency, but could improve the quality of the identified knowledge, and the quality of the way the external knowledge is used. Further, a sourcing expert may not have contextual knowledge regarding the problem that motivated the search, which would be needed to identify the most relevant solution (Cohen and Levinthal 1990; Tyre and von Hippel 1997). Difficulty explaining an uncertain problem may also limit the quality of the answer; perhaps team members do not fully understand their problem until they implement a search themselves (Fleming 2001).

When a team has a broad sourcing pattern, it also limits its reliance on knowledge transfer from the sourcing expert to the other team members. When the actual user of the external knowledge is involved in bringing the knowledge into the team, she may recognize that more of the knowledge artifact applies to her present situation (or even other situations) than a gatekeeper could realize (Cohen and Levinthal 1990) and so is likely to transfer more of the knowledge (Staats, Brunner and Upton 2011). Further, direct interaction with useful knowledge increases an individual's understanding of the salience of that knowledge, as compared with receiving a summary or translation from a colleague (Gino 2008).

Finally, broad sourcing may change the way that the knowledge is eventually applied. As many or all team members engage in information processing, they are more likely to adapt and transform the knowledge to their specific need. This is consistent with the information processing perspective that finds that decentralized team communications result in more innovative output. When more team members are actively sourcing knowledge, there is greater opportunity for feedback, reflection, and error correction, all of which may improve the quality of the output (Argyris and Schön 1978; Tushman 1979). Organization design scholars also argue that decentralized knowledge generation may improve the quality of the solution (Burns and Stalker 1961).

Broad knowledge sourcing may also improve the quality of the integrated whole. Work *quality* is an interdependent project requires an integrated, holistic view of the unique output and the ability to flawlessly combine individual components (Faraj and Sproull 2000). Mistakes in any one area can severely compromise an entire project. Because interdependent outcomes like quality require greater information processing capability (Tushman 1979), and more broad knowledge sourcing may improve the information processing capability of the team, more broad knowledge sourcing is likely to be related to project quality. Thus, we hypothesize:

> HYPOTHESIS 2: *Specialized team knowledge sourcing is negatively associated with project quality performance.*

2.2 Organizing Team Knowledge Sourcing and Knowledge Sharing

In the previous section, we argue that different knowledge sourcing patterns will lead to performance tradeoffs. Next we consider conditions under which a specialized or broad sourcing pattern will be even more strongly associated with team performance. In particular, we consider how conditions that impede knowledge sharing *within* the team may amplify the consequences of how the team brings knowledge *into* the team (i.e., sources knowledge). We examine the moderating influence of three team compositional factors that have been shown to complicate within-team knowledge sharing – geographic dispersion, lack of team familiarity, and subgroup strength – on the relationships between sourcing pattern and different

performance outcomes (Gibson and Vermeulen 2003; Huckman et al. 2009; O'Leary and Mortensen 2010).

2.2.1 Knowledge Sharing: Team Geographic Dispersion

Teams with geographically dispersed members can access and integrate knowledge and expertise from each other around the globe without meeting face-to-face, offering attractive cost structures (Jarvenpaa and Leidner 1999; Gibson and Gibbs 2006; O'Leary and Cummings 2007). However, such dispersed teams face obvious barriers to communication. Their members often must rely on electronic communication media, which can decrease the volume and quality of communication (Maznevski and Chudoba 2000) and increase conflict (Hinds and Mortensen 2005). Dispersed teams also have a difficult time knowing who knows what within a team (Cramton 2001). Both the communication difficulties and the lack of mutual knowledge significantly impede knowledge sharing within globally dispersed work teams (Cramton 2001; Polzer et al. 2006).

These difficulties may be particularly problematic for teams that have adopted a specialized sourcing pattern. Specialized sourcing depends on effective knowledge sharing – the team member who needs the knowledge must effectively communicate his or her focal problem to the person who will search for the knowledge (Griffith, Sawyer and Neale 2003). The sourcing expert must then search, correctly identify the best knowledge for the problem, and return and transfer the identified knowledge to the inquiring team member. Each step of this search and transfer problem will be complicated by the lack of face-to-face communication and resulting process difficulties commonly seen in dispersed teams. Communication of the focal problem is likely to be incomplete or unclear because geographically team members tend to focus too much on their common knowledge when interactions take place (Stasser and Titus 1987; Larson et al. 1998). Poorly specifying the problem will make it more difficult to identify the best knowledge during the search. Once knowledge has been sourced, the transfer to the end-user may prove more difficult for geographically dispersed teams (Griffith et al. 2003), especially because electronically mediated communication allows for less rich knowledge transfer (Maznevski and Chudoba 2000; Majchrzak, Malhotra and John 2005; Boh 2008). We therefore hypothesize that geographic

dispersion will interact with concentration to result in worse performance for both efficiency and quality performance:

> HYPOTHESIS 3:   *Specialized team knowledge sourcing and team geographic dispersion have a negative interaction effect on efficiency and quality performance.*

2.2.2 Knowledge Sharing: Lack of Team Familiarity

Team familiarity – prior shared work experience – has been shown to positively affect project team performance (Reagans, Argote and Brooks 2005; Espinosa et al. 2007; Huckman et al. 2009). Teams that lack familiarity do not perform as well as familiar teams. Previous research has demonstrated that team familiarity facilitates knowledge sharing; familiar teams are better able to locate specialized knowledge within the team (Wegner 1987; Littlepage, Robison and Reddington 1997), and team familiarity also helps teams develop shared language and mutual understanding (Arrow 1974; Cramton 2001; Staats 2011). Familiarity in teams is also associated with increased trust, positive social acceptance, and psychological safety, all of which support effective knowledge sharing (Gruenfeld et al. 1996; Uzzi 1997; Edmondson 1999). These dynamics, typically associated with familiarity, also support the capability to integrate knowledge (Gardner, Gino and Staats 2011). Teams that lack the social benefits of familiarity may be less able to share knowledge in the team effectively.

The knowledge-sharing benefits associated with team familiarity may be particularly valuable for teams using specialized knowledge sourcing. The communication advantages of shared language, mutual knowledge, and shared context increase the likelihood that a sourcing expert will identify the most relevant organizational knowledge for the current problem. Transfer of the knowledge will also be more efficient and rich in familiar teams. The positive shared beliefs held by familiar teams such as psychological safety, trust, and positive social acceptance, will also facilitate knowledge transfer within the team (Stasser and Titus 1987; Larson et al. 1998; Huckman and Staats 2011). Teams made up of members who do not have shared history working together will have a more difficult time with the

knowledge sharing required by a specialized sourcing pattern. As a result we hypothesize:

> HYPOTHESIS 4:   *Specialized team knowledge sourcing and team familiarity have a positive interaction effect on efficiency and quality performance.*

2.2.3 Knowledge Sharing: Subgroup Strength

A third condition that has been shown to complicate knowledge sharing in teams is having strong subgroups, which can lead to fault lines. Fault lines are "hypothetical dividing lines that may split a group intro subgroups based on one or more attributes (Lau and Murnighan 1998: 328)." Subgroups complicate knowledge sharing because they heighten the salience of differences between team members (Earley and Mosakowski 2000). Individuals may view members of their own subgroup as in-group members, while viewing those in other subgroups as part of an out-group, which creates biases as to how the individual perceives information from each (Allport 1954). Strong subgroups created by overlapping identity groups may lead to higher levels of conflict and lower levels of trust, both which can hinder knowledge sharing (Lau and Murnighan 1998; Li and Hambrick 2005).

The fault lines between subgroups function as boundaries that inhibit knowledge sharing. Gibson and Vermeulen (2003) showed that a knowledge management initiative benefited team learning but that the existence of subgroups lessened this relationship. Lau and Murnighan's (2005) experimental study suggested one possible reason for this. They found that within-group communication aided team effectiveness for groups with weak subgroups, but were of no benefit for groups with strong subgroups. The knowledge sharing difficulties for teams with strong subgroups will be particularly problematic for teams with specialized knowledge sourcing, which rely on knowledge sharing to turn sourced knowledge into better team performance. Thus, we hypothesize:

> HYPOTHESIS 5:   *Specialized team knowledge sourcing and subgroup strength have a negative interaction effect on efficiency and quality performance.*

3. Setting, Data, and Empirical Strategy

3.1 Setting

To test our hypotheses, we study the relationship between team knowledge sourcing patterns and team performance in software system development projects at Wipro Technologies. Software projects are challenging to coordinate and to deliver successfully (Faraj and Sproull 2000; Huckman et al. 2009). A development project includes identifying customer requirements, creating a custom solution, writing software code that fulfills the requirements, and testing the code (Boehm 1981). Development projects are ideal for our purposes because objective performance measures and control variables are available for all development teams, allowing for comparisons across projects.

Our data set covers software projects completed during 2008 and 2009. During this time period the company was actively competing with both leading Indian (e.g., TCS and Infosys) and Western (e.g., IBM and Accenture) firms. The leadership at Wipro believed that ongoing success centered on executing projects efficiently and with high quality. These executives considered identifying, storing, and making available existing organizational knowledge essential for project success. They also believed that the ability of employees to identify and exploit valuable knowledge was at risk because of the increasing number of employees, the geographic distribution of the organization, and the ongoing turnover of engineers. Building on their award-winning knowledge management strategy (Wipro 2007), the company in 2007 committed additional resources to improve its knowledge repository in an effort to support knowledge sourcing within the organization. Wipro changed its knowledge management website (named KNet), and as part of the changes included advanced analytic technology to permit detailed tracking of repository use. Employees were encouraged to upload and download content and content was screened to ensure high quality. Any employee could use the system. In addition to inviting input from the workforce, a knowledge management team also developed content. Note that although there is some reusable code in the knowledge repository, this is a small percentage of the knowledge artifacts; most artifacts are documents. Although the company actively encouraged use of the repository during 2008-2009, they did

not have a specific policy on how the teams should use it and instead teams were left to make decisions on their own as to how to structure knowledge sourcing.

We also conducted interviews to develop deeper insight into the dynamics and performance effects of team knowledge sourcing patterns. During three visits to Wipro's operations in India we interviewed over thirty individuals involved in the knowledge management initiative, including senior executives at Wipro, and project team leaders and members. We draw on these data in our discussion to add texture and to illustrate mechanisms and practices that may underlie our results. We do not conduct formal analyses of these data, but rather offer them to illuminate some of the ways that knowledge sourcing in Wipro teams unfolds. For example, describing benefits of sourcing knowledge from the knowledge repository, one project manager noted, "There are many different types of documents on KNet. For example, there are case studies of prior projects that talk about benefits, problems, customer value, new innovations, and best practices. There are also documents that explain how a specific aspect of a technology or domain works. These include details about solving particular problems such as the flow of development, steps to follow, and examples. Also, there are reusable forms where someone can post code or an object." Another project manager added, "Early in a project, we can go to KNet and use it to find best practices. We can look at case studies and see lessons learned and what issues different projects faced. All of this helps the team deliver better."

3.2 Data

The empirical analysis draws on three sources of data: knowledge sourcing from the knowledge repository, project outcomes and characteristics, and human capital information. With the restructuring of its knowledge repository in 2007, the company gained ability to track person-level use – the number of unique downloads (i.e., from different URLs) a person makes on a day. This data consists of person-day observations from January 1, 2008, to December 31, 2009. Each URL corresponds to a unique knowledge artifact in the knowledge repository. Although we know the number of unique downloads each person made per day, we have no information on the content a user views or how long they view an artifact.

Our second data set consists of information about the 460 development projects that started on or

after January 1, 2008, and finished by December 31, 2009. We restrict our analysis to these projects as we wish to examine only those projects for which we have complete data on knowledge sourcing. The number of projects in our final analysis is reduced further by 130 projects because we drop projects that were the only project for a given customer to run our model which controls for customer effects. We control for customer effects in order to account for time-invariant aspects of customers that could affect our performance measures (e.g., customer processes and legacy systems, difficulty in gaining access to systems, etc.) Finally, we also collect complete human capital information on the 9,554 individuals involved in these projects. This information includes both demographic information and individual project assignments since 2000.

The median project team is our analysis includes 20 individuals. Team members usually work on only one project at a time. Teams are not assembled to complete multiple projects together, but rather after a project is finished the team breaks up and moves on to other projects. While some individuals on a team may have worked together previously, the number varies dramatically between teams (see the team familiarity variable, below for a measure of team members' prior interactions). Inside of a project team the project manager is the appointed leader. The other two roles within the team are middle managers, who both write code and manage a smaller sub-team, and project engineers, who write code.

3.2.1 Dependent Variables. We use two classes of dependent variables: effort deviation, to measure efficiency, and post-delivery defects, to measure quality. Table 1 includes summary statistics.

*****************************INSERT TABLE 1 HERE*****************************

Project Efficiency. To evaluate the efficiency of a project, we examine whether the project meets its estimated number of hours of work to be completed. We construct a variable, *effort deviation*, by subtracting the project's estimated effort (in person-hours) from its actual effort required (in person-hours), then dividing that difference by the estimated effort to normalize for project size. Effort estimates are initially made by sales and pre-sales personnel at Wipro. The estimate may be changed during the course of a project, usually as a result of the customer's changing a project's scope. To make sure estimates are not altered for inappropriate reasons (e.g., because a project is falling behind), a project

manager must receive both customer and internal Wipro management approval for the change. We use the revised estimates to calculate effort deviation, because these estimates most accurately capture the final goals and objectives of a project.

Project Quality. In addition to examining project efficiency, we also examine post-delivery defects as a measure for project quality. After the completion of a development project, customer acceptance testing takes place in most, but not all projects. During customer acceptance testing, the customer or a third party tests the code against the project's pre-specified metrics. The output of this process is a count of the number of defects, or *post-delivery defects,* a commonly used quality metric in software engineering (Boehm 1981). As a process check, in cases where zero defects are recorded, the internal auditing group confirms testing took place.

3.2.2 Independent Variables. Our hypotheses focus on the concentration of team knowledge sourcing and the interactions with geographic dispersion, team familiarity, and subgroup strength. We standardize each of these variables, by subtracting the mean and dividing by the standard deviation to aid interpretation and limit multicollinearity of interaction effects (Aiken and West 1991).

Specialized team knowledge sourcing. To examine the specialization of team knowledge sourcing, we utilize the Herfindahl measure, commonly used to capture the distribution of a characteristic across team members (Harrison and Klein 2007). The measure is calculated by first computing the percentage of downloads made by each individual on the team (i.e., each individual's downloads / total team downloads), then squaring and summing these values for the entire team. Thus, in a project with highly specialized use (e.g., only one person downloaded from the knowledge repository), the Herfindahl measure would equal one. Alternatively, for a project with equal use across all team members, the measure would equal 1 / N, where N equals the size of the team.

Geographic dispersion. Project teams in our context deploy individuals across *two* locations: Wipro's Indian facilities – "offshore" – and clients' offices – "onsite" (31% of projects locate individuals at only one site). During this time period, management reported that teams working onsite or offshore were typically collocated in the same building on the same floor. The majority of project work was completed

for US and European clients and so teams across the two sites were temporally separated as well (although customer identifiers in the data are anonymized so we do not have data on where a customer is located).

Therefore, we first calculate the percentage of the team's hours spent offshore. However, because we are interested in how work is dispersed rather than how much of the work is completed offshore, we then create a variable, *geographic dispersion,* equal to the offshore percentage if less than half the work is completed offshore, or equal to one minus the offshore percentage if more than half the work is completed offshore. We calculate the variable in this manner because a team that is 65% offshore and 35% onsite is as equally dispersed as a team that is 35% offshore and 65% onsite. We use hours of effort, rather than team members' location, to construct the variable because hours more precisely measures the distribution of project effort. However, substituting a variable constructed by averaging the location of each team member (coded one or zero for whether an individual is offshore or onsite) yields similar results for all hypotheses.[2]

Team familiarity. We operationalize team familiarity in line with prior work (Reagans et al. 2005; Espinosa et al. 2007; Huckman et al. 2009). We calculate the team familiarity variable by summing the number of times each unique dyad on the team has worked together on any project over the prior three years. We divide that sum by the number of unique dyads on the team, to scale the variable (Reagans et al. 2005). We use a three-year window to capture the fact that as with other types of knowledge (e.g., Argote, Beckman and Epple 1990) team familiarity may decay over time. In addition, the three-year window capture multiple project cycles, since the average project lasts approximately seven months. We conduct sensitivity analyses with two- and four-year windows and generate the same pattern of results.

Subgroup strength. We calculate subgroup strength using a similar approach to Gibson and Vermeulen (2003). Their measure is calculated based on gender, ethnic background, functional background, team tenure, and age. We do not include functional background because all team members in our setting are

---

[2] In other words, we assign each individual a "1" if she is located offshore and a "0" if she is onsite. Then we average all values for the entire team. We then create a variable equal to the average if the average is less than ½ or equal to one minus the average if it is greater than ½ (for the same logic as with geographic dispersion).

software engineers. Additionally, we do not include ethnic background because that data is not available from the company. According to our contacts at Wipro, the vast majority of team members are Indian workers educated at Indian universities. Finally, we substitute firm tenure for team tenure because in our context teams are not ongoing entities, but rather are broken up at the end of each project (Huckman et al. 2009).

Thus using gender, age, and firm tenure we first calculate the overlap between each member of the team on each dimension. For gender the overlap is one (same gender) or zero. For firm tenure overlap is calculated by dividing the smaller firm tenure value by the larger firm tenure value, thus bounding the number between zero and one. The same approach is used for age except we adjust for the fact that age is bounded between 20 and 54. With these values we then calculate subgroup strength by calculating the standard deviation across the three overlap measures for all team pairs (Gibson and Vermeulen 2003).

3.2.3 Control Variables.

As detailed in Table 2, we control for a number of variables, including the mean level of knowledge sourcing by the team, which may affect team performance.

*****************************INSERT TABLE 2 HERE*****************************

3.3 Empirical Strategy

Because effort deviation is a continuous variable and post-delivery defects is a count variable, we use different regression models for each. In each case, we wish to control for time-invariant attributes of customers that can affect performance (Greene 2003). For effort deviation, we use a fixed-effects linear regression model because a Hausman test rejects the null hypothesis that the random effects model is consistent ($p<0.05$). For post-delivery defects, we use a conditional fixed effects negative binomial model, as the data exhibits overdispersion and a Hausman test again rejects the null hypothesis that the random effects model is consistent (Cameron and Trivedi 1998). Since the model conditions on the customer, as opposed to just including a customer fixed effect, the model drops all customers in which the defect variable does not vary from zero. This yields 255 projects for our defect analysis. We compare the projects used in the quality models to the seventy-five projects used only in the efficiency models and

find no statistically significant differences in the independent variables of interest between the two.

Given that our study examines individuals nested in teams a natural consideration is to use hierarchical linear modeling (HLM). However, while our independent variables capture variance at the individual- and team-level our dependent variable only captures variance at the team-level. In such situations, as noted by Hofmann (2002), since "only between-group variance in the independent variable can predict variance in the dependent variable (p. 263)," regression analyses at the team-level are both appropriate and necessary. Therefore, we structure our analyses as detailed above.

4. Results

Table 3 shows results for the study's regression models. Columns 1, 2, and 3 correspond to the models for effort deviation, while Columns 4, 5, and 6 report results for post-delivery defects. Column 1 reports the model without the knowledge sourcing variables, which are included in Column 2. As Hypothesis 1 predicts, we find that higher specialized sourcing is related to lower effort deviation (i.e., improved project efficiency). An increase of 0.26 in specialized team knowledge sourcing (a one-standard-deviation increase) is related to a 42% decrease in average effort deviation. We also note that the amount of team knowledge sourcing is also related to improved project efficiency. An increase of 3.55 in team knowledge sourcing (a one-standard-deviation increase) is related to an 8% decrease in average effort deviation.

*****************************INSERT TABLE 3 HERE****************************

In Column 3, we add the interaction variables. Although Hypothesis 3 predicts that the interaction of specialized team knowledge sourcing and team geographic dispersion is related to worse project efficiency, the coefficient on the interaction terms is not statistically significant, failing to provide support for the hypothesis. Examining the interaction of specialized team knowledge sourcing and team familiarity, we find the coefficient is negative and statistically significant. This result supports Hypothesis 4. Finally, the coefficient on the interaction of specialized team knowledge sourcing and subgroup strength is negative and statistically significant, not supporting Hypothesis 5.

In Figure 1, we investigate the two statistically significant interaction terms in more detail by

plotting the effects for high and low values of specialized team knowledge sourcing (one standard deviation above and below the mean) and team familiarity (one standard deviation above the mean and no prior familiarity) / subgroup strength (one standard deviation above and below the mean). For team familiarity, we find that at high values of team familiarity, high specialized team knowledge sourcing is related to significantly better efficiency performance. At low values of team familiarity low and high specialized team knowledge sourcing do not show a significant difference (i.e., the simple slope is not significant). Figure 1 shows the same pattern for subgroup strength – more specialized team knowledge sourcing is related to significantly better efficiency performance than low specialized team knowledge sourcing for high values of subgroup strength and the difference between high and low specialized knowledge sourcing is not significant at low values of subgroup strength. We explore this surprising subgroup result in more detail in the discussion section.

****************************INSERT FIGURES 1a and 1b HERE****************************

In Column 4 of Table 3, we report the defect model without knowledge sourcing variables, which we include in Column 5. In support of Hypothesis 2 we see that more specialized knowledge repository use is related to worse performance (i.e., higher expected defects). An increase of 0.26 in specialized team knowledge sourcing (a one-standard-deviation increase) is related to a 40% increase in expected defects. In Column 6, we add the interaction terms. Consistent with Hypotheses 4 and 6 we find that the interaction of specialized team knowledge sourcing with both geographic dispersion and subgroup strength are positive and statistically significant (related to more expected defects). However, the interaction of specialized team knowledge sourcing and team familiarity is not statistically significant, failing to support Hypothesis 5. Figure 2 plots the statistically significant results for low and high values of the variables and provides graphical support for Hypotheses 4 and 6. In both plots the simple slope for the high values of the moderator (geographic dispersion and subgroup strength) are statistically significant while the simple slopes for the low values of the moderator are not. Thus, more specialized team knowledge sourcing is related to worse quality performance for both high geographic dispersion and high subgroup strength.

*************************INSERT FIGURES 2a and 2b HERE*************************

In our research model we specify geographic dispersion, team familiarity, and subgroup strength as moderators of the relationship between specialized team knowledge sourcing and team performance. Although these hypothesized relationships are supported by prior theory, an additional question is whether these variables might lead to different patterns of emergent team organization for knowledge sourcing. We run the first stage of a mediation model to investigate whether geographic dispersion, team familiarity, or subgroup strength predict specialized team knowledge sourcing (Baron and Kenny 1986). None of the variables are statistically significant, increasing our confidence that our model is correctly specified.

Additionally, we conduct two checks for multicollinearity issues. First, we estimate models using ordinary least squares to calculate variance inflation factors (VIF). All VIFs are less than three, which is less than the suggested threshold of ten (Cohen et al. 2003). Second, we rerun the models with interaction terms with each interaction added individually and the results are consistent, giving us confidence in our results.

5. Discussion

Teams that carry out complex project work need access to accurate, timely knowledge to perform well. Many firms support ambitious knowledge management strategies to fill that need: AMR Research estimated that in 2007 U.S. companies spent approximately $73 billion on knowledge repository and collaboration technologies (Murphy and Hackbush 2007). The present study shows that knowledge management presents a paradox for teams; they cannot "have it all" when sourcing knowledge. Although having everyone on the team interface with the knowledge system does appear to enhance quality performance, it can harm efficiency performance. Our research also shows that these effects are moderated by three conditions affecting knowledge sharing within the team: geographic dispersion, team familiarity, and subgroup strength.

Our interviews provided multiple perspectives of participants at Wipro, consistently supporting

our idea that teams adopt different sourcing patterns. The interviews also allowed us to ensure that the differences in approach we identified were not dictated by Wipro managers or policies; teams were free to work out how to interact with the knowledge repository. Further, we gained deeper insight into how sourcing patterns work within software teams.

Several project managers revealed awareness of what we have labeled a specialized sourcing pattern. For example, one said, "I have been on projects where two or three people will become formal experts on KM [knowledge management]. People know that these experts really know the system. This can help improve project performance since they [team members] can get the right knowledge very quickly. …So if someone has a question, they may go to the expert to ask about structuring a query or more commonly, they may just ask the expert to get the information for them." In follow-up questioning the manager clarified that individuals were informal experts; their roles were not formally assigned, but rather emerged in the process of the team's work. Several other managers offered benefits of what we call a broad sourcing pattern. For example, one manager noted, "Finding a solution through KM can quicken the pace. However, sometimes in searching for the solution, you can examine the process the other person used to solve the problem. If you see the process yourself, then you not only understand the solution better, but you could also improvise or find a better way." A team member explained, "Unless you go and look yourself you may not find the answer [to a specific question]. I can look in one area and find nothing, but I may then get an idea or find something completely different that solves my problem."

Examining what factors give rise to the knowledge sourcing patterns we identified is an important avenue for future research. Perhaps more experienced project leaders are able to scan the environment and organize knowledge sourcing within the team to match the needs of the environment. Such actions would be of theoretical importance, but would not invalidate the relevance of our findings – we are investigating the more proximal relationship between knowledge sourcing and team performance. However, in understanding what factors lead to different patterns of organization for work it will be possible to improve theoretical models and practical guidance (Brown and Eisenhardt 1995).

One surprising result warrants further discussion. As expected, the interaction of subgroup

strength and concentration of team knowledge sourcing is related to worse quality performance. But subgroup strength and specialized sourcing are related to *better* efficiency performance. This result might be due to a similar dynamic to that found by Lau and Murnighan (2005), who showed that improved within-subgroup interactions compensated for the worse across-subgroup interactions across a divisible task (Gibson and Vermeulen 2003). It is possible that within-subgroup knowledge sharing is enough to overcome across-subgroup knowledge sharing challenges, for project efficiency. Software work can be subdivided and a project may still be delivered efficiently, albeit with potential quality issues due to integration challenges. As noted by Li and Hambrick (2005), the subgroup concept is a valuable theoretical one, but how it plays out in practice is quite complicated. There is a need and opportunity for future work to study the differences between within-subgroup knowledge sharing and across-subgroup knowledge sharing.

5.1 Limitations and Future Work

Several limitations to our work must be noted. First, although our dataset provides detailed information on when and how many times individuals source knowledge from the repository, we lack detail on what knowledge they access. Future work should explore how similarities and differences in the knowledge sourced by team members affect performance. Second, our use of only a linear term for subgroup strength is consistent with some work on subgroups (Lau and Murnighan 2005; Li and Hambrick 2005), but many other studies use a quadratic term to model the effect of subgroups (Gibson and Vermeulen 2003). Adding a quadratic term to our model does not generate statistically significant results. It is possible that our study examines a different range of subgroups than work that finds quadratic effects (although the summary statistics suggest that ranges are comparable, with ours being slightly higher, on average) or that our lack of variance on ethnic background fails to activate curvilinear effects. Alternatively, it is possible that the effects of subgroups may not be curvilinear with respect to team performance or with archival data. Future work should explore when curvilinear effects do and do not appear.

Third, our data on the geographic location of individuals only measures whether individuals on a

team are located at Wipro's facilities in India or at the customer's location, not whether they are spread out across either. Wipro managers reported that during this time teams were typically located in one place both in India and at the customer. The level of detail we have on individuals' locations permits us to construct a continuous variable for geographic dispersion at the team level and is thus, more specific than may prior studies that code dispersion with a one or a zero. Although it would be preferable to have the additional information, our dispersion variable remains an informative measure of one condition that inhibits the free flow of knowledge sharing.

Fourth, we examine different ways that teams organize knowledge sourcing, but note that other types of variation in team sourcing will also likely affect performance in different ways. Future work could explore whether sourcing early vs. late in the project has a differential affect. Additionally, future research could also examine differences in the individuals who serve as gatekeepers and whether certain individual characteristics increase an individual's likelihood to succeed in the role. Fifth, we find that different patterns of knowledge sourcing are relate to project performance within one type of project (software development projects). Future work should explore these effects within other contexts and project types (e.g., radical innovation projects). Finally, although the empirical dataset we use is both large and detailed, it comes from one organization. This is a necessary, but less than ideal, consequence of gaining access to a research site and collecting such data. Future work should examine these findings in other companies and settings.

5.2 Managerial Implications

Our findings have at least two important implications for managers. The first is that teams face a performance trade-off when it comes to making use of the knowledge assets held within the firm. Therefore, helping a team recognize this trade-off and deliberately choose a knowledge sourcing approach that matches performance requirements may increase the likelihood of successful performance on that metric. Second, within-team knowledge sharing amplifies the difficulty of specialized knowledge sourcing, in some cases. Managers may be able to use these findings to compose teams that are better able to take advantage of certain knowledge sourcing structures or to support teams likely to have difficulty

sharing knowledge.

5.3 Conclusion

In this research we make several contributions to the literature on teams and knowledge management. First, we adapt the construct of knowledge sourcing (Gray and Meister 2004) to the team-level of analysis, and consider how team approaches to knowledge sourcing can vary. This allows us to demonstrate the complexities of team knowledge sourcing and to explicate a performance paradox for teams sourcing knowledge from outside the team. Second, we demonstrate the value of a compositional approach to team knowledge sourcing (Kozlowski et al. 1999; Harrison and Klein 2007) by showing that there are predictable performance differences based on how team members source knowledge, even when holding the amount of team knowledge sourcing constant. Third, we use data on actual system use to test and support our theory. These extensive archival data allow us to contribute both research and practice. Prior research on knowledge repositories and knowledge sourcing has been limited to survey data, but Devaraj and Kohli (2003) demonstrated that the perceived and actual use of an IT system may not match (Straub, Limayem and Karahanna-Evaristo 1995). Therefore, there is great value to testing hypotheses using archival data. This also provides valuable information for managers. A senior manager at our field site lamented, "We believe that our knowledge management initiative has value, but we have no empirical evidence to support that view." We are able to show, using archival data, that knowledge sourcing from the knowledge repository is related to improved performance. Finally, we respond to calls to investigate the relationship between knowledge sourcing and knowledge sharing (Thomas-Hunt, Ogden and Neale 2003; Gray and Meister 2004). In examining how within-team knowledge sharing moderates the team knowledge sourcing and performance relationship we contribute to research on geographically-dispersed teams, team familiarity, and subgroup strength. Altogether in this paper our findings offer insight for theory and practice into how team organization, organizational knowledge resources, and within-team knowledge sharing can aid team performance.

6. References

Aiken, L. S. and S. G. West (1991). <u>Multiple Regression: Testing and Interpreting Interactions</u>. Thousand Oaks, CA, Sage Publications.

Alavi, M. and D. E. Leidner (2001). "Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues." <u>MIS Quarterly</u> **25**(1): 107-136.

Allen, T. J., D. M. S. Lee and M. L. Tushman (1980). "R & D performance as a function of internal communication, project management, and the nature of the work." <u>IEEE Transactions on Engineering Management</u> **27**: 2-12.

Allen, T. J., M. L. Tushman and D. M. S. Lee (1979). "Technology transfer as a function of position in the spectrum from research through development to technical services." <u>Academy of Management Journal</u> **22**(4): 694-708.

Allport, G. W. (1954). <u>The Nature of Prejudice</u>. Cambridge, MA, Addison-Wesley.

Ancona, D. and H. Bresman (2007). <u>X-teams: How to Build Teams That Lead, Innovate and Succeed</u>. Boston, Harvard Business School Press.

Argote, L., S. L. Beckman and D. Epple (1990). "The persistence and transfer of learning in industrial settings." <u>Management Science</u> **36**(2): 140-154.

Argyris, C. and D. A. Schön (1978). <u>Organizational learning</u>. Reading, Mass., Addison-Wesley Pub. Co.

Arrow, K. J. (1974). <u>The Limits of Organization</u>. New York, W.W. Norton & Co.

Banker, R. D. and C. F. Kemerer (1989). "Scale economies in new software development." <u>IEEE Transactions on Software Engineering</u> **15**(10): 1199-1206.

Baron, R. M. and D. A. Kenny (1986). "The mediator–moderator variable distinction in social psychological research, conceptual, strategic, and statistical considerations." <u>Journal of Personality and Social Psychology</u> **51**(6): 1173–1182.

Bock, G. W., R. Sabherwal and Z. J. Qian (2008). "The effect of social context on the success of knowledge repository systems." <u>IEEE Transactions on Engineering Management</u> **55**(4): 536-551.

Bock, G. W., R. W. Zmud, Y. G. Kim and J. N. Lee (2005). "Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate." <u>MIS Quarterly</u> **29**(1): 87-111.

Boehm, B. (1981). <u>Software Engineering Economics</u>. New York, Prentice Hall.

Boh, W. F. (2008). "Reuse of knowledge assets from repositories: A mixed methods study." <u>Information & Management</u> **45**(6): 365-375.

Brown, S. L. and K. M. Eisenhardt (1995). "Product development: Past research, present findings, and future directions." <u>Academy of Management Review</u> **20**(2): 343-378.

Burns, T. and G. M. Stalker (1961). <u>The Management of Innovation</u>. London, Tavistock Publications.

Cabrera, A. and E. F. Cabrera (2002). "Knowledge-sharing dilemmas." <u>Organization Studies</u> **23**(5): 687-710.

Cameron, A. C. and P. K. Trivedi (1998). <u>Regression Analysis of Count Data</u>. New York, Cambridge University Press.

Cohen, J., P. Cohen, L. S. Aiken and S. G. West (2003). <u>Applied Multiple Regression—Correlation Analysis for the Behavioral Sciences, 3rd ed</u>. Mahwah, NJ., Lawrence Erlbaum.

Cohen, W. M. and D. A. Levinthal (1990). "Absorptive capacity: A new perspective on learning and innovation." <u>Administrative Science Quarterly</u> **35**(1): 128-152.

Cramton, C. D. (2001). "The mutual knowledge problem and its consequences for dispersed collaboration." <u>Organization Science</u> **12**(3): 346.

Cummings, J. N. (2004). "Work groups, structural diversity, and knowledge sharing in a global organization." <u>Management Science</u> **50**(3): 352–364.

Daft, R. L. and K. E. Weick (1984). "Toward a model of organizations as interpretation systems." <u>Academy of Management Review</u> **9**(2): 284-295.

Davenport, T. H. and L. Prusak (1998). <u>Working Knowledge: How Organizations Manage What They Know</u>. Boston, Harvard Business School Press.

Devaraj, S. and R. Kohli (2003). "Performance impacts of information technology: Is actual usage the missing link?" <u>Management Science</u> **49**(3): 273-289.

Earley, P. C. and E. Mosakowski (2000). "Creating hybrid team cultures: An empirical test of transnational team functioning." <u>Academy of Management Journal</u> **43**(1): 26-49.

Edmondson, A. (1999). "Psychological safety and learning behavior in work teams." <u>Administrative Science Quarterly</u> **44**(2): 350-383.

Edmondson, A. C. and I. A. Nembhard (2009). "Product development and learning in project teams: The challenges are the benefits." <u>Journal of Product Innovation Management</u> **26**(2): 123-138.

Espinosa, J. A., S. A. Slaughter, R. E. Kraut and J. D. Herbsleb (2007). "Familiarity, complexity, and team performance in geographically distributed software development." <u>Organization Science</u> **18**(4): 613-630.

Ethiraj, S. K., P. Kale, M. S. Krishnan and J. V. Singh (2005). "Where do capabilities come from and how do they matter? A study in the software services industry." <u>Strategic Management Journal</u> **26**(1): 25-45.

Faraj, S. and L. Sproull (2000). "Coordinating expertise in software development teams." <u>Management Science</u> **46**(12): 1554-1568.

Fleming, L. (2001). "Recombinant uncertainty in technological search." <u>Management Science</u> **47**(1): 117-132.

Fleming, L. and O. Sorenson (2004). "Science as a map in technological search." <u>Strategic Management Journal</u> **25**(8-9): 909-928.

Gardner, H. K., F. Gino and B. R. Staats (Forthcoming). "Dynamically integrating knowledge in teams: Transforming resources into performance." Academy of Management Journal.

Gibson, C. and F. Vermeulen (2003). "A healthy divide: Subgroups as a stimulus for team learning behavior." Administrative Science Quarterly **48**(2): 202-239.

Gibson, C. B. and J. 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): 451-495.

Gino, F. (2008). "Do we listen to advice just because we paid for it? The impact of advice cost on its use." Organizational Behavior and Human Decision Processes **107**(2): 234-245.

Gray, P. H. and A. Durcikova (2005). "The role of knowledge repositories in technical support environments: Speed versus learning in user performance." Journal of Management Information Systems **22**(3): 159-190.

Gray, P. H. and D. B. Meister (2004). "Knowledge sourcing effectiveness." Management Science **50**(6): 821-834.

Greene, W. H. (2003). Econometric Analysis. Upper Saddle River, N.J., Prentice Hall.

Griffith, T. L., J. E. Sawyer and M. A. Neale (2003). "Virtualness and knowledge in teams: Managing the love triangle of organizations, individuals, and information technology." MIS Quarterly **27**(2): 265-287.

Gruenfeld, D. H., E. A. Mannix, K. Y. Williams and M. A. Neale (1996). "Group composition and decision making: How member familiarity and information distribution affect process and performance." Organizational Behavior and Human Decision Processes **67**(1): 1-15.

Haas, M. R. (2006). "Knowledge gathering, team capabilities, and project performance in challenging work environments." Management Science **52**(8): 1170-1184.

Hackman, J. R. (2002). Leading Teams: Setting the Stage for Great Performances. Boston, HBS Press.

Hackman, J. R. and N. Katz (2010). Group behavior and performance. Handbook of Social Psychology. S. T. Fiske, D. T. Gilbert and G. Lindzey. New York, Wiley. **5th ed**.

Hansen, M. T. and M. R. Haas (2001). "Competing for attention in knowledge markets: Electronic document dissemination in a management consulting company." Administrative Science Quarterly **46**(1): 1-28.

Hansen, M. T., N. Nohria and T. Tierney (1999). "What's your strategy for managing knowledge?" Harvard Business Review **77**(2): 106-116.

Harrison, D. A. and K. J. Klein (2007). "What's the difference? Diversity constructs as separation, variety, or disparity in organizations." Academy of Management Review **32**(4): 1199-1228.

Heath, C. and N. Staudenmayer (2000). Coordination neglect: How lay theories of organizing complicate coordination in organizations. Research in Organizational Behavior. B. M. Staw and R. I. Sutton.

Henderson, R. M. and K. B. Clark (1990). "Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms." Administrative Science Quarterly **35**(1): 9-30.

Hinds, P. J. and M. Mortensen (2005). "Understanding conflict in geographically distributed teams: The moderating effects of shared identity, shared context, and spontaneous communication." Organization Science **16**(3): 290-307.

Hofmann, D. A. (2002). Issues in multilevel research: Theory development, measurement, and analysis. Handbook of Research Methods in Industrial and Organizational Psychology. S. G. Rogelberg. Oxford, Blackwell**:** 247-274.

Huckman, R. S. and B. 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): 310-328.

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

Jarvenpaa, S. L. and D. E. Leidner (1999). "Communication and trust in global virtual teams." Organization Science **10**(6): 791-815.

Katz, R. and M. Tushman (1979). "Communication patterns, project performance, and task characteristics: An empirical evaluation and integration in an R&D setting." Organizational Behavior and Human Performance **23**(2): 139-162.

Kozlowski, S. W. J., S. M. Gully, E. R. Nason and E. M. Smith (1999). Developing adaptive teams: A theory of compilation and performance across levels and time. The Changing Nature of Performance: Implications for Staffing, Motivation, and Development. D. R. Ilgen and E. D. Pulakos. San Francisco, Jossey-Bass**:** 240-292.

Larson, J. R., Jr. (2010). In Search of Synergy in Small Group Performance. New York, Psychology Press.

Larson, J. R., Jr., C. Christensen, T. M. Franz and A. S. Abbott (1998). "Diagnosing groups: The pooling, management, and impact of shared and unshared case information in team-based medical decision making." Journal of Personality and Social Psychology **75**(1): 93-108.

Lau, D. C. and J. K. Murnighan (1998). "Demographic diversity and faultlines: The compositional dynamics of organizational groups." Academy of Management Review **23**(2): 325-340.

Lau, D. C. and J. K. Murnighan (2005). "Interactions with groups and subgroups: The effects of demographic faultlines." Academy of Management Journal **48**(4): 645-659.

Li, J. and D. C. Hambrick (2005). "Factional groups: A new vantage on demographic faultlines, conflict, and disintegration in work teams." Academy of Management Journal **48**(5): 794-813.

Littlepage, G., W. Robison and K. Reddington (1997). "Effects of task experience and group experience on group performance, member ability, and recognition of expertise." Organizational Behavior and Human Decision Processes **69**(2): 133-147.

MacCormack, A., R. Verganti and M. Iansiti (2001). "Developing products on "Internet time": The anatomy of a flexible development process." Management Science **47**(1): 133-150.

Majchrzak, A., L. P. Cooper and O. E. Neece (2004). "Knowledge reuse for innovation." Management Science **50**(2): 174-188.

Majchrzak, A., A. Malhotra and R. John (2005). "Perceived individual collaboration know-how development through information technology–enabled contextualization." Information Systems Research **16**(1): 9 - 27.

March, J. G. (1991). "Exploration and exploitation in organizational learning." Organization Science **2**(1): 71-87.

Maznevski, M. L. and K. M. Chudoba (2000). "Bridging space over time: Global virtual team dynamics and effectiveness." Organization Science **11**(5): 473-492.

Murphy, J. and J. Hackbush. (2007). "The knowledge management spending report, 2007–2008." AMR Research Retrieved August 17, 2010.

Nickerson, J. A. and T. R. Zenger (2004). "A knowledge-based theory of the firm: The problem-solving perspective." Organization Science **15**(6): 617-632.

O'Leary, M. B. and J. N. Cummings (2007). "The spatial, temporal, and configurational characteristics of geographic dispersion in teams." MIS Quarterly **31**(3): 433-452.

O'Leary, M. B. and M. Mortensen (2010). "Go (con)figure: Subgroups, imbalance, and isolates in geographically dispersed teams." Organization Science **21**(1): 115-131.

Polzer, J. T., C. B. Crisp, S. L. Jarvenpaa and J. W. Kim (2006). "Extending the faultline model to geographically dispersed teams: How colocated subgroups can impair group functioning." Academy of Management Journal **49**(4): 679-692.

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

Simon, H. A. (1962). "The architecture of complexity." Proceedings of the American Philosophical Society **106**(6): 467-482.

Staats, B. R. (Forthcoming). "Unpacking team familiarity: The effect of geographic location and hierarchical role." Production and Operations Management.

Staats, B. R., D. J. Brunner and D. M. Upton (2011). "Lean principles, learning, and knowledge work: Evidence from a software services provider." Journal of Operations Management **29**(5): 376-390.

Stasser, G. and W. Titus (1987). "Effects of information load and percentage of shared information on the dissemination of unshared information during group discussion." Journal of Personality and Social Psychology **53**(1): 81-93.

Stein, E. W. and V. Zwass (1995). "Actualizing organizational memory with information systems." Information Systems Research **6**(2): 85-117.

Straub, D., M. Limayem and E. Karahanna-Evaristo (1995). "Measuring system usage: Implications for IS theory testing " Management Science **41**(8): 1328-1342.

Szulanski, G. (1996). "Exploring internal stickiness: Impediments to the transfer of best practice within the firm." Strategic Management Journal **17**(Special Issue: Knowledge and the Firm): 27-43.

Szulanski, G. (2000). "The process of knowledge transfer: A diachronic analysis of stickiness." Organizational Behavior and Human Decision Processes **82**(1): 9-27.

Thomas-Hunt, M. C., T. Y. Ogden and M. A. Neale (2003). "Who's really sharing? Effects of social and expert status on knowledge exchange within groups." Management Science **49**(4): 464-477.

Tushman, M. L. (1977). "Special boundary roles in the innovation process." Administrative Science Quarterly **22**(4): 587-605.

Tushman, M. L. (1979). "Work characteristics and subunit communication structure: A contingency analysis." Administrative Science Quarterly **24**(1): 82-98.

Tushman, M. L. and R. Katz (1980). "External communication and project performance: An investigation into the role of gatekeepers." Management Science **26**(11): 1071-1085.

Tyre, M. J. and E. von Hippel (1997). "The situated nature of adaptive learning in organizations." Organization Science **8**(1): 71-83.

Uzzi, B. (1997). "Social structure and competition in interfirm networks: The paradox of embeddedness." Administrative Science Quarterly **42**(1): 35-67.

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. Theories of Group Behavior. G. Mullen and G. Goethals. New York, Springer-Verlag**:** 185-208.

7. Figures and Tables
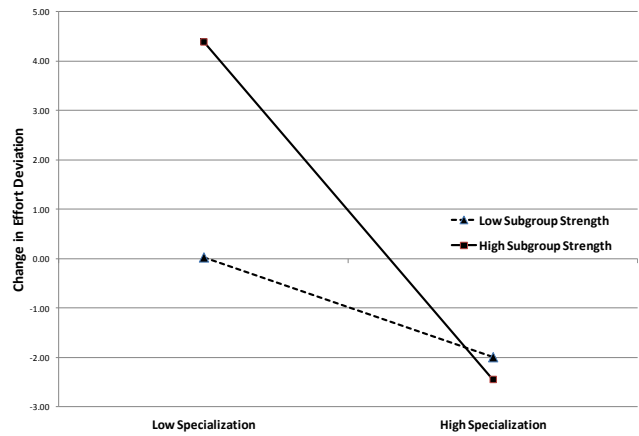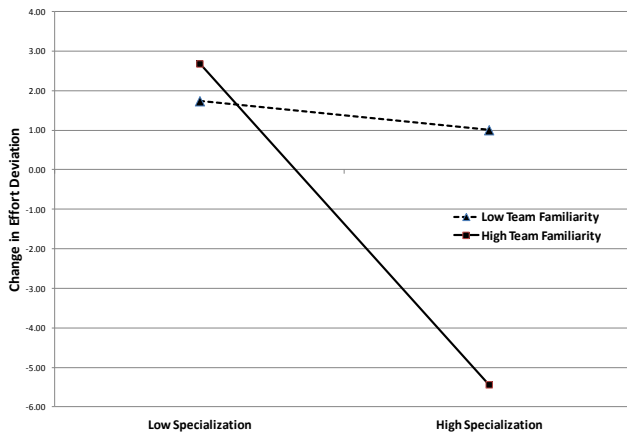
a. Specialized team knowledge sourcing and team familiarity. **b.** Specialized team knowledge sourcing and subgroup strength.

Figure 1. Plots of the interaction effects on effort deviation.
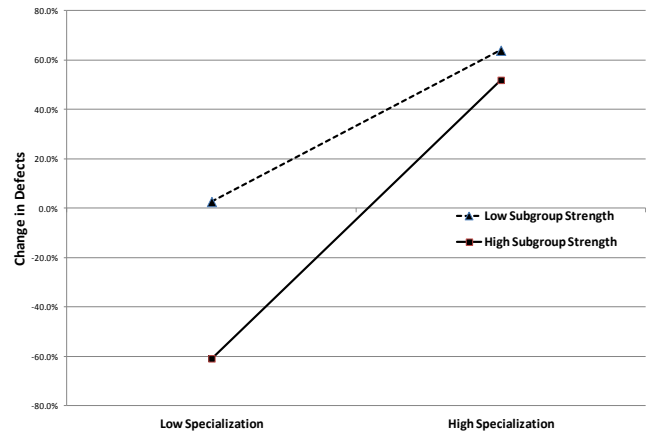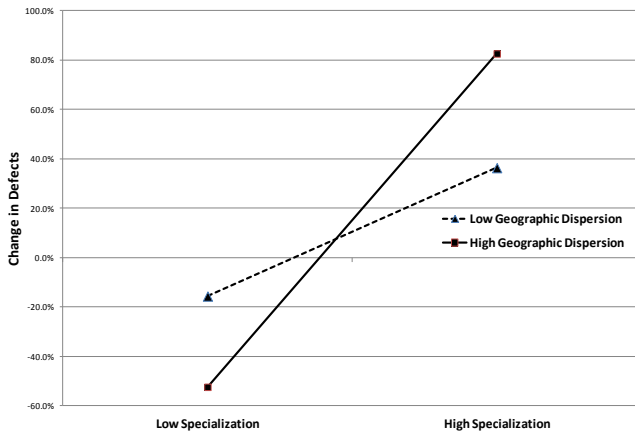
a. Specialized team knowledge sourcing and geographic dispersion. **b.** Specialized team knowledge sourcing and subgroup strength.

Figure 2. Plots of the interaction effects on defects.

Table 1. Summary statistics and correlation table of variables of interest
(*n*= 330, except for post-delivery defects, where *n* = 255).

Variable	Mean	σ	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Effort Deviation	-5.18	10.91													
2. Defects	13.25	73.26	0.06												
3. Team Knowledge Sourcing	5.66	3.55	0.04	0.05											
4. Specialized Team Knowledge Sourcing[a]	0.24	0.26	**-0.18**	-0.06	**-0.15**										
5. Team Heterogeneity	1.13	0.10	0.05	-0.04	0.00	**-0.14**									
6. Subgroup Strength[a]	0.75	0.18	0.08	0.11	**0.12**	**-0.37**	**0.30**								
7. Geographic Dispersion[a]	0.14	0.14	**-0.19**	0.03	-0.09	-0.07	0.04	0.01							
8. Team Familiarity[a]	0.39	0.77	-0.10	-0.07	0.03	-0.05	**-0.20**	**-0.12**	0.12						
9. Role Experience	1.64	0.80	-0.02	0.04	0.01	-0.06	-0.07	0.02	0.04	**0.17**					
10. Project Scale	3.15	1.87	0.05	0.00	**0.18**	**-0.15**	0.10	**0.11**	-0.06	-0.04	0.01				
11. Estimated Effort	8.86	1.00	0.10	**0.21**	**0.18**	**-0.42**	0.11	**0.42**	**0.12**	**-0.19**	-0.03	**0.26**			
12. Team Size	3.05	0.80	0.11	**0.17**	0.05	**-0.48**	0.10	**0.71**	0.06	-0.01	0.03	**0.15**	**0.55**		
13. Estimated Duration	5.45	0.60	**0.20**	**0.13**	**0.28**	**-0.27**	0.08	**0.30**	0.07	**-0.21**	**-0.18**	**0.19**	**0.62**	**0.34**	
14. Contract Type	0.46	0.50	**-0.25**	-0.06	**-0.12**	0.09	0.02	-0.06	0.04	-0.10	0.05	0.00	0.05	**-0.13**	-0.04

Note. Bold denotes significance of less than 5%.
[a] In models this variable is standardized by subtracting the mean and dividing by the standard deviation. Values here are before standardization.

Table 2. Control variables included in empirical models.

Team knowledge sourcing	We measure a team's knowledge sourcing from the repository by counting the number of times a team member accesses a unique URL on any given day during the duration of a project, summing this total for all team members, then dividing by the number of team members. We calculate the mean to normalize for the effects of team size.
Team heterogeneity	To calculate our measure of team heterogeneity we use the same three overlap measures as in subgroup strength. We sum the overlap between team members and divide by the number of pairs on the team, yielding a measure of homogeneity. We then take the inverse of the measure to yield team heterogeneity (Gibson and Vermeulen 2003).
Role experience	Prior work has established that role experience is a useful measure for team experience (Huckman et al. 2009). Role experience averages the years individual on the team has served in their present, hierarchical role (i.e. project manager, middle manager, or project engineer). To calculate the team value, we weight each individual's value by the number of days he or she was on the team. Calculating the variable without the weights does not change our results in a substantive manner.
Project scale	We measure project scale using kilolines of new source code (KLOC). KLOC is a commonly used measure to evaluate software scale (MacCormack, Verganti and Iansiti 2001). The company uses a regimented approach for counting lines of code. As software has been shown to exhibit scale effects, we use the log of KLOC in our models (Banker and Kemerer 1989).
Estimated effort	Projects involving more hours of effort may in turn be more difficult to complete. Therefore, we include the log of the estimated total person-hours for a given project. We use the estimated value because a project that is delivered over budget (i.e., takes more hours) would have a larger actual effort value than a project that meets its estimates.
Team size	If a team becomes too large, then adding members could create coordination or integration challenges (Heath and Staudenmayer 2000). Alternatively, if a team is small, then adding team members could prove useful, as doing so increases the knowledge resources available to the team (Hackman 2002). Therefore, we control for team size by including the log of total personnel who worked on the project.
Estimated duration	We control for project duration because a longer project may be more difficult or face greater employee attrition (Ethiraj et al. 2005) than a shorter project. We use the log of the estimated value (in days) to avoid the same endogeniety concern as with effort. Estimated duration and estimated effort are correlated (Pearson coefficient = 0.62). We include both variables as they capture some separate information about a project. However, dropping either variable generates the same pattern of results that we present.
Contract type	Development contracts use either a time-and-materials (i.e., cost-plus) structure or a fixed-price structure. In the former case, a customer pays the negotiated rate for number of hours worked on the project while in the latter case a set payment is agreed to prior to the start of the project. We use an indicator variable, contract type, to control for these differences, such that the variable equals one if the contract is for a fixed price, and zero if it is for time and materials.
Software languages, number and type	We control for the number and type(s) of software languages used in a project. For the former, we include an indicator set to one if a project uses more than one software language (which is the case for 53% of projects). For the latter, we include indicator variables for the different languages used (C, C++, Java, query, markup, BASIC).
Technologies	Projects may include multiple technology classes (e.g., client server, e-commerce). Therefore, we include an indicator equal to one if a project includes more than one technology, and zero otherwise (10% and 90% of projects, respectively).

Table 3. Summary results of the regression of effort deviation and post-delivery defects on specialized team knowledge sourcing (n = 330 and 255, respectively).

	Dep Variable: Effort Deviation			Dep Variable: Defects		
	(1)	(2)	(3)	(4)	(5)	(6)
Team Knowledge Sourcing		-0.438**	-0.442**		-0.012	-0.011
		(0.183)	(0.174)		(0.029)	(0.029)
Specialized Team Knowledge Sourcing[a]		-2.180***	-2.213***		0.339***	0.456***
		(0.747)	(0.725)		(0.118)	(0.115)
Specialized Team Knowledge Sourcing × Geographic Dispersion			-0.845			0.216**
			(0.552)			(0.098)
Specialized Team Knowledge Sourcing × Team Familiarity			-3.693***			0.086
			(0.738)			(0.129)
Specialized Team Knowledge Sourcing × Subgroup Strength			-1.203**			0.222**
			(0.549)			(0.113)
Geographic Dispersion[a]	-1.742**	-2.045***	-1.851**	-0.048	0.006	-0.070
	(0.775)	(0.761)	(0.731)	(0.108)	(0.108)	(0.112)
Team Familiarity[a]	-1.546**	-1.460**	-2.749***	-0.337**	-0.283*	-0.380**
	(0.752)	(0.735)	(0.751)	(0.166)	(0.159)	(0.188)
Subgroup Strength[a]	-0.028	0.235	0.978	-0.074	-0.042	-0.261
	(0.945)	(0.925)	(0.999)	(0.144)	(0.142)	(0.168)
Team Heterogeneity	-4.983	-8.811	-8.948	0.479	0.910	1.104
	(7.012)	(6.921)	(6.593)	(1.133)	(1.147)	(1.156)
Role Experience	0.495	0.492	0.921	-0.020	0.002	-0.045
	(0.823)	(0.807)	(0.775)	(0.122)	(0.121)	(0.123)
Project Scale	-0.181	-0.036	0.117	0.065	0.073	0.059
	(0.356)	(0.353)	(0.336)	(0.063)	(0.063)	(0.066)
Estimated Effort	0.039	-0.150	-0.383	0.109	0.138	0.246
	(0.956)	(0.941)	(0.906)	(0.146)	(0.147)	(0.158)
Team Size	-0.928	-2.413*	-2.141	-0.055	0.074	0.103
	(1.324)	(1.353)	(1.316)	(0.206)	(0.214)	(0.219)
Estimated Duration	3.709**	4.456***	4.131***	-0.178	-0.074	-0.114
	(1.475)	(1.496)	(1.420)	(0.246)	(0.240)	(0.237)
Contract Type	-2.670	-2.685	-2.523	0.294	0.219	0.129
	(1.721)	(1.681)	(1.596)	(0.232)	(0.236)	(0.241)
Constant	-15.486	-7.300	-5.912	-2.341	-3.975**	-4.757***
	(11.069)	(11.344)	(10.814)	(1.681)	(1.749)	(1.793)
Observations	330	330	330	255	255	255
Overall R^2	0.1214	0.1699	0.2680	-	-	-
Log-likelihood	-	-	-	-377.6520	-373.9234	-368.5806
F Statistic	2.0634**	2.6956***	4.0788***	-	-	-
Wald chi-squared	-	-	-	14.5503	23.4635	34.4002**

Notes. *, ** and *** denote signficance at the 5%, 1% and 0.1% levels, respectively. Effort deviation models are GLS fixed-effects models with heteroskedasticity robust standard errors, clustered on the customer. Defect models are conditional fixed effects negative binomial regression models that condition on the customer. All models include, but results are not shown for the following variables: number of languages, start year, software language, and number of technologies.

[a] Variable is standardized by subtracting the mean and dividing by the standard deviation.