6-2015

Which Problems to Solve? Online Knowledge Sharing and Attention Allocation in Organizations

Martine R. HAAS
University of Pennsylvania

Paola CRISCULO
Imperial College London

Gerard GEORGE
Singapore Management University, ggeorge@smu.edu.sg

Citation

HAAS, Martine R.; CRISCULO, Paola; and GEORGE, Gerard. Which Problems to Solve? Online Knowledge Sharing and Attention Allocation in Organizations. (2015). *Academy of Management Journal*. 58, (3), 680-711. Research Collection Lee Kong Chian School Of Business.
Available at: https://ink.library.smu.edu.sg/lkcsb_research/4620

Published in Academy of Management Journal, 2015, 58(3), pp. 680-711.
https://doi.org/10.5465/amj.2013.0263

WHICH PROBLEMS TO SOLVE? ONLINE KNOWLEDGE SHARING AND ATTENTION ALLOCATION IN ORGANIZATIONS

MARTINE R. HAAS*
University of Pennsylvania

PAOLA CRISCUOLO*
Imperial College London

GERARD GEORGE
Singapore Management University

Why do individuals allocate attention to specific problems in organizations? Viewing online knowledge sharing as a matching process between knowledge providers and problems, we examine attention allocation in the context of an online community within which knowledge providers respond to problems posted by other organization members. We argue that knowledge providers are more likely to allocate attention to solving problems that more closely match their expertise, but that decisions to allocate attention are also influenced by problem characteristics such as length, breadth, and novelty, as well as by problem crowding. Analyzing 1,251 realized matches and 12,510 nonrealized matches among knowledge providers and problems posted over a 32-month period on an online discussion forum within a global engineering firm, we find evidence to support our claim that attention allocation is driven by the features of a particular provider–problem match, thereby shifting the discourse from knowledge provider–seeker relationships to knowledge provider–problem matches. The implications for theories of knowledge sharing, matching processes, and managerial attention are discussed.

In the digital economy, individuals and organizations are awash with information. With more than 3.2 billion social networking users, 3.9 billion active e-mail users, and 400 million tweets a day, the rise of social media has generated vast amounts of information content. Businesses own more than 900 million mailboxes worldwide, which account for more than 100 billion work-related e-mails sent and received daily (Radicati, 2013), with the average manager spending 28% of his or her workday sending and answering e-mails (McKinsey Global Institute, 2012). One report on social technologies, defined as "information technology (IT) products and services that enable the formation and operation of online communities, where participants have distributed access to content and distributed rights to create, add and/or modify content" (McKinsey Global Institute, 2012: 1), estimates that a value of over US$1 trillion can be realized annually through social technologies and that individual employee productivity can be enhanced by 20–25%. This explosion of social technologies has the power to transform organizations and organizational life.

* The first two authors contributed equally to this paper. Please direct correspondence to Martine Haas, mrhaas@wharton.upenn.edu

The authors thank Jason Colquitt and the three anonymous AMJ reviewers, as well as Ammon Salter, Bill McEvily, Henrich Greve, Chiara Criscuolo, and Anne Ter Wal, for their helpful comments and advice. Martine Haas also thanks participants in research seminars at Stanford University (Department of Management Science and Engineering), Boston University (School of Management), Harvard Business School, INSEAD, University of Virginia (Darden), and Dartmouth (Tuck) for their valuable feedback. Paola Criscuolo gratefully acknowledges the support of the British Engineering and Physical Sciences Research Council (grant ref. EP/F036930/1). Gerry George acknowledges the support of the Economic and Social Research Council Professorial Fellowship (grant ref. RES-051-27-0321).

Within organizations, the critical processes of learning, innovation, and performance increasingly depend on how members of the organization utilize such social technologies to share knowledge (Argote, McEvily, & Reagans, 2003; Brown & Duguid, 2002; Sambamurthy & Subramani, 2005). To facilitate knowledge sharing, many large organizations have established electronic communities of practice and introduced social technology platforms to support them, such as online discussion forums (or message boards) on which employees can post problems related to their work and share solutions with each other.[1] Such platforms are potentially valuable for knowledge sharing, but their proliferation can contribute to an increasing sense of information overload among employees (Davenport & Beck, 2001; Dean & Webb, 2011). In a world of information overload, attention becomes a critical scarce resource (Simon, 1947). Accordingly, the finite attention of employees becomes a key constraint on problem solving (Cyert & March, 1963; March & Simon, 1958; Ocasio, 1997). Faced with a growing number of problems seeking solutions via social technology platforms, individuals who might be able to provide solutions to others' problems must decide not only whether to allocate attention to offering solutions at all, but also which problems to address. Since information overload is a growing challenge, the question of why organization members decide to allocate attention to addressing particular problems online is an increasingly urgent concern for organizations.

To explain why individuals choose to respond to problems online at all, prior research on online knowledge sharing in organizations has pointed toward social motives such as reputation enhancement, commitment to the community, and generalized reciprocity (e.g., Chiu, Hsu, & Wang, 2006; Constant, Sproull, & Kiesler, 1996; Wasko & Faraj, 2000). Such benefits are particularly important because many organizations do not offer explicit rewards or incentives for online knowledge sharing among their employees. However, individuals are likely to be concerned about the costs, as well as the benefits, of spending time and effort responding to others' problems online, since attention is a finite resource. Moreover, the question of *which* problems individuals choose to address online has not been addressed in prior studies. Applying findings from research on interpersonal knowledge sharing to the online context suggests that individuals might be more likely to respond to problems from other individuals with whom they have connections based on factors such as social similarity, physical proximity, or prior familiarity (e.g., Espinosa, Slaughter, Kraut, & Herbsleb, 2007; Quigley, Tesluk, Locke, & Bartol, 2007; Reagans, 2011). Yet in online settings individuals often respond to problems posted by others with whom they have no such connections, to the extent that Constant and colleagues (1996) noted that online knowledge sharing seems to be driven by "the kindness of strangers." The implication is that individuals may choose to respond to problems for reasons beyond interpersonal connections—perhaps for reasons related to the problem itself.

In this study, we examine knowledge sharing in the context of an intraorganizational online discussion forum: a social technology platform that provides an informal setting in which knowledge seekers (that is, employees who are searching for solutions to problems) can post task-related questions and knowledge providers (that is, employees who can offer solutions to those problems) can post answers. We explore why knowledge providers allocate attention to some problems rather than others in this context by shifting perspectives to focus on provider–problem matching rather than provider–seeker relationships, and by taking into account the costs, as well as the benefits, that these providers can expect to incur. The context is of theoretical relevance for our research question because there are many problems seeking solutions, and individuals decide which problems to address, if any. It is also of practical importance since many large, dispersed organizations use online discussion forums to facilitate knowledge sharing among their employees (e.g., Davenport & Prusak, 2000; Kane & Alavi, 2007). Other social technology platforms, such as e-mail, document repositories, and groupware, are widely used for knowledge sharing within firms too (e.g., Ahuja & Carley, 1999; Bock, Zmud, Kim, & Lee, 2005; Kankanhalli, Tan, & Wei, 2005), but the distinctive advantages of an online discussion forum are that knowledge seekers can search both broadly and efficiently for solutions to their problems, and can obtain immediate, customized responses from knowledge providers whom they might not otherwise reach.

In order for knowledge seekers to receive responses, however, knowledge providers have to

[1] Online discussion forums are also increasingly used to share knowledge across and outside organizational borders (e.g., Faraj, Jarvenpaa, & Majchrzak, 2011; Jeppesen & Lakhani, 2010).

decide to allocate attention to addressing their problems. Attention allocation involves the focusing of time and effort on a stimulus (James, 1890; Kahneman, 1973). While attention allocation can be mindful or less mindful (Levinthal & Rerup, 2006; Weick & Sutcliffe, 2006), we focus on deliberate decisions to allocate attention to solving particular problems, as manifested by whether or not an individual posts a response to a problem in an online discussion forum. We draw on organizational theories of matching processes (e.g., Mitsuhashi & Greve, 2009; Vissa, 2011) to analyze why individuals allocate attention to some problems rather than others. As a baseline, we propose that this matching process will be influenced by how closely the expertise possessed by the knowledge provider matches the expertise required by the problem. We then consider the effect of other problem characteristics that can attract attention, but which also create cognitive load for a knowledge provider, such as the problem's length, breadth, and novelty, as well as the effects of problem crowding in the form of concurrently posted problems that can attract attention to the forum, but which also compete with the focal problem for attention. Finally, we propose that expertise matching can moderate the effects of problem characteristics and problem crowding on a provider's decision to allocate attention to a problem by increasing the benefits of attention allocation and reducing the costs created by cognitive load and competitive crowding. We test our hypotheses using field data from a global engineering firm, in which organization members utilized an online discussion forum to post problems and share solutions related to structural engineering—a core competence of the firm.

Our study contributes to conversations about how attention is allocated inside organizations, with broader implications for information processing in social technology environments. By viewing attention allocation as a matching process, we bring matching theory into the organization, and highlight the theoretical importance of matches between particular knowledge providers and particular problems for influencing what receives attention. Our attention perspective sheds light on how knowledge sharing is shaped by factors that influence the costs, as well as the benefits, that providers can incur when allocating attention to problems, while controlling for provider–seeker relationships and other factors that may influence this activity. Examining the increasingly pressing question of how attention is allocated in the context of an online discussion forum also contributes to our understanding of online knowledge sharing, a phenomenon of growing practical significance within and across organizations. Perhaps, at its core, the study helps to illuminate a challenge of central importance to organizations: understanding why some problems get solved, while others do not.

AN ATTENTION PERSPECTIVE ON KNOWLEDGE SHARING

Knowledge Sharing in an Online Community

In large, dispersed organizations in which knowledge is widely distributed, online communities often utilize social technology platforms, such as discussion forums, to enable knowledge seekers to access solutions to problems from knowledge providers across the organization. By posting questions to an online discussion forum, individuals can search beyond their own social networks, minimize coordination costs, and receive answers from others whom they did not know could offer them. For an online discussion forum to function effectively, however, voluntary participation from knowledge providers is necessary.

Prior research on online knowledge sharing has identified a variety of social motivations that may lead knowledge providers to contribute solutions to problems. For example, in an early study of advice giving in a technical online community, Constant and colleagues (1996) found that the benefits to knowledge providers in a Fortune 100 company seemed to arise primarily from the gratification of helping colleagues and from the reputational enhancement gained by demonstrating expertise. Wasko and Faraj (2005) found that members of a legal professional association were more likely to contribute to an online discussion forum if they felt that they had more to share, anticipated reputational benefits, and were structurally embedded in the professional network. Chiu and colleagues (2006) found that perceived norms of reciprocity, as well as social ties and community identification, increased the propensity to share knowledge in a professional IT network in Taiwan. Other studies have found evidence for effects of functional role and hierarchical status (Ahuja, Galletta, & Carley, 2003), user experience and recognition (Jeppesen & Frederiksen, 2006), perceived identity verification (Ma & Agarwal, 2007), and self-efficacy (Hsu, Ju, Yen, & Chang, 2007). In the related context of electronic document repositories, scholars have uncovered additional factors that can affect contributions, ranging from individual self-worth to generalized trust, a climate of fairness, and organizational rewards (e.g., Bock et al., 2005; Kankanhalli et al., 2005). Taken together, these studies suggest that a range of motives lead individuals to contribute solutions to problems.

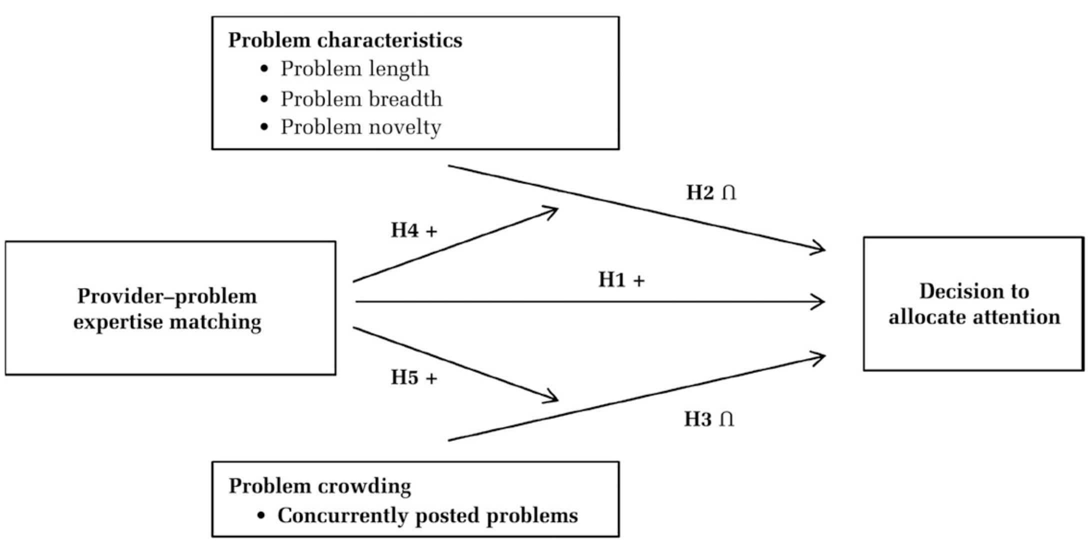
However, much of this research examines general propensities to contribute, rather than why specific contributions are made. To the extent that previous research on online knowledge sharing has focused on dyadic exchanges rather than overall contributions, it has assumed that contributions depend on relationships between providers and seekers (e.g., Constant et al., 1996). This focus on provider–seeker relationships builds on research on knowledge sharing through personal networks, which has shown that providers are often more willing to share knowledge with seekers to whom they feel personally connected. A personal connection between a provider and a seeker may arise from social similarity or homophily, which encourages interaction between individuals with similar demographic or social characteristics (e.g., McPherson, Smith-Lovin, & Cook, 2001; Reagans, 2005). It may come from physical proximity, which exposes individuals to each other and makes it easy for them to access each other (e.g., Allen, 1977; Cummings, 2004). It also may come from prior familiarity, which establishes mutual knowledge and expectations of ongoing reciprocity (e.g., Cramton, 2001; Espinosa et al., 2007). Yet research on online knowledge sharing has shown that contributions often occur in the absence of interpersonal homophily, proximity, or familiarity (e.g., Constant et al., 1996). Moreover, there is mixed evidence for the importance of reciprocity in online knowledge sharing, with some arguing that expectations of reciprocity matter (e.g., Chiu et al., 2006; Walther, Anderson, & Park, 1994), while others find that they do not (e.g., Constant et al., 1996; Wasko & Faraj, 2005). The implication is that knowledge sharing in an online discussion forum is driven by factors beyond provider–seeker relationships based on homophily, proximity, familiarity, and reciprocity.

To shed new light on what these factors might be, we take a different perspective from prior research by viewing knowledge sharing in an online discussion forum as driven by provider–problem matching, rather than by provider–seeker relationships. In our empirics, we account for the likelihood that a provider contributes to the forum at all as a precondition for our analyses, and we also control for provider–seeker relationships, but our theoretical arguments and our main empirical analyses focus specifically on a provider's decision to allocate attention to a particular problem. Below, we argue that this decision is driven by the expertise match between the provider and the problem, as well as by other problem characteristics and by problem crowding. Our hypotheses are summarized in Figure 1.

Provider–Problem Expertise Matching

Prior research has shown that knowledge providers' levels of expertise are important in influencing their

FIGURE 1
Model of Provider–Problem Attention Allocation in an Online Discussion Forum



overall contributions to an online community. For example, Constant et al. (1996) found that individuals with higher levels of expertise were more likely to contribute answers to an online discussion forum, and Wasko and Faraj (2000) confirmed that individuals were less likely to contribute answers when they felt that their expertise was inadequate. These studies relied on self-reported levels of expertise and overall contributions, and did not examine either the content of the providers' expertise or the content of the problems to be addressed. Nevertheless, they suggest that a provider who can offer expertise that more closely matches the expertise required by a focal problem will be more likely to decide to allocate attention to that problem.

In part, providers whose expertise more closely matches the expertise required by a problem may see greater benefits in allocating attention to that problem. These benefits may arise from the satisfaction of helping others (e.g., Dudley & Cortina, 2008) or from the value that they anticipate creating (Nahapiet & Ghoshal, 1998), as well as from the prospect of enhancing their reputation or encouraging future reciprocity by using their expertise to provide a good solution (e.g., Chiu et al., 2006; Constant et al., 1996; Wasko & Faraj, 2000). To the extent that such benefits are anticipated, they can be expected to be greater when there is a closer match between the content of provider's expertise and the content of the problem.

Additionally, providers whose expertise matches a focal problem also face lower costs of attention allocation. Closer expertise matching increases the likelihood that the provider has the absorptive capacity necessary to understand the problem (Cohen & Levinthal, 1990; Zahra & George, 2002). Further, evidence from geographically distributed teams suggests that overlapping expertise enhances mutual knowledge, reducing the costs of understanding and responding to others' problems (Kotha, George, & Srikanth, 2013). This makes it quicker and easier for the provider to make sense of the problem, to grasp its intricacies, contingencies, and ramifications, to situate it in a broader knowledge landscape, and to identify and articulate a solution (Sole & Edmondson, 2002; Thomas, Sussman, & Henderson, 2001; Tsai, 2001). In contrast, potential knowledge providers may find it more difficult to solve, or even to understand, a problem when the content of their expertise and the content of the problem are more divergent, owing to their lower absorptive capacity and the insularity of their knowledge base (George, Kotha, & Zheng, 2008), making it more costly to respond effectively.

Because a potential knowledge provider with expertise that more closely matches a problem can anticipate both greater benefits and lower costs from allocating attention to that problem, we expect that a closer expertise match will increase the likelihood that a provider allocates attention to a problem in an online discussion forum. Thus, as a baseline prediction, we expect:

Hypothesis 1. The likelihood that a provider allocates attention to a focal problem will be positively related to the closeness of the provider–problem expertise match.

Problem Characteristics

Theories of selective attention suggest that a problem is likely to attract attention if it is salient—that is, if it stands out more relative to alternative targets for attention allocation (McArthur, 1981; Taylor & Fiske, 1978). According to cognitive psychologists, salience does not rely on prior preferences for a particular kind of stimulus; instead, attention is drawn selectively to a stimulus on exposure (Higgins, 1996). The implication is that the characteristics of a problem that make it more salient can increase the likelihood that a potential knowledge provider will allocate attention to that problem.

In an online discussion forum, those characteristics of a problem that can make it more salient for a potential knowledge provider include its length, breadth, and novelty. Longer problems take up more space on the screen, dominating a provider's field of vision and crowding out other stimuli (see Parkhurst, Law, & Niebur, 2002; Wolfe & Horowitz, 2004). In addition, once the potential knowledge provider starts to read, longer problems may also be more engaging, as the provider gets drawn into the details of the situation presented in the problem. Broader problems are more likely to touch on a domain of expertise of interest to a provider, offering a hook that captures the provider's attention (see Cohen, March, & Olsen, 1972; March 1994). For example, a problem that mentions tennis, soccer, and baseball is more likely to attract attention than one that mentions only tennis. Again, once the potential knowledge provider starts to read, broader problems may also be more engaging as they connect the domain of expertise that originally attracted the provider to other domains of expertise. More novel problems can attract attention as a result of distinctiveness effects (e.g., Gardner, 1985; Nelson, 1979; Taylor & Fiske, 1978). Prior research has shown that executives pay more attention to issues that subordinates portray as more novel (Dutton, Ashford, O'Neill, & Lawrence, 2001), and unfamiliar terrains are more likely to capture search attention during new product development (Li, Maggitti, Smith, Tesluk, & Katila, 2013). Similarly, in an online discussion forum, problems that are novel relative to other problems are likely to stand out more, and thus to attract attention. Additionally, research on open-source software suggests that tackling novel problems can be intrinsically motivating, as well as extrinsically rewarding, since solving them can serve as a reputation-enhancing signal to the community (Lakhani & Wolf, 2005).

Some length, breadth, and novelty thus can help to attract attention to a problem. At high levels, however, length, breadth, and novelty can impose costs on potential knowledge providers that may reduce their propensity to allocate attention to the problem. When a problem is high in length, breadth, or novelty, it creates cognitive load for a provider, in the form of nontrivial information-processing demands (Sweller, 1988). This cognitive load may result from intrinsic or extraneous factors: intrinsic cognitive load is generated by the problem's inherent level of difficulty, while extraneous load is generated by the way in which it is presented—for example describing a square in writing imposes more extraneous load than providing a picture of a square (Chandler & Sweller, 1991; Paas, van Gog, & Sweller, 2010).

Long problems may generate greater intrinsic load than short problems because they are inherently more difficult, requiring a potential knowledge provider to utilize more complex cognitive schemas to process them. They may also create greater extraneous load because they are wordier than necessary, perhaps because of an unfocused portrayal of the problem or an unnecessarily detailed description, such that the provider needs to expend greater effort to understand the question, to distil the essential information, and to articulate a response. Broad problems are likely to generate greater cognitive load than narrow problems because they are likely to be inherently more difficult, controlling for expertise matching, since they span multiple domains of expertise, and thus require that providers construct and utilize more complex schemas to process them (Dane, 2010). Problems that are novel for the forum likewise are likely to create greater cognitive load than problems that are routine for the forum, either because they are inherently more difficult to address if the knowledge provider does not possess the schemas needed to process them and must build them from scratch, or because even if the provider does possess the necessary schemas, the novelty of the problem for the forum means that there is still more work to be done to help others to understand the solution than is needed for a routine problem (George et al., 2008; Kotha et al., 2013). Thus problems with high levels of length, breadth, or novelty impose higher cognitive loads on potential knowledge providers, increasing the costs of allocating attention to such problems. The consequence, as Kahneman (1973: 53) pointed out, can be that "excessively complex stimuli are treated as irrelevant noise and no longer attract attention."

In summary, a problem that is very short, narrow, or routine can fail to attract attention from a potential knowledge provider as a result of low salience, while a problem that is very long, broad, or novel can be offputting because of high cognitive load. Taken together, these arguments suggest that there will be a curvilinear relationship between a problem's length, breadth, or novelty and a potential knowledge provider's decision to allocate attention to that problem, such that greater length, breadth, and novelty will have positive effects on the likelihood of attention allocation, but only up to a point, after which greater length, breadth, and novelty will have negative effects on the likelihood of attention allocation. Hence we propose:

> *Hypothesis 2. The likelihood that a provider allocates attention to a focal problem will be curvilinearly related to the problem's (a) length, (b) breadth, and (c) novelty, in an inverse U shape.*

Problem Crowding

While cognitive psychologists' theories of selective attention suggest that the characteristics of a problem itself are important for attention allocation, organizational scholars have proposed theories of selective attention tailored to organizational settings that suggest that the extent to which problems attract attention will vary not only with the characteristics of those problems themselves, but also with the contexts within which the problems are embedded (e.g., March & Olsen, 1976; Weick, 1979). Ocasio (1997) calls this "the principle of situated attention"—that is, what individuals focus on depends on the particular situation or context in which they find themselves (Ross & Nisbett, 1991).

In an online community's discussion forum, a central feature of the context that can be expected to influence the allocation of attention to a focal problem is problem crowding, in the form of the number of other problems concurrently posted to the forum (see Jones, David, & Rafaeli, 2004; Piezunka & Dahlander, 2014). Theories of intra-organizational ecology and competition for attention suggest that the full set of problems that are seeking solutions can influence how individuals allocate attention to specific problems within firms (e.g., Burgelman, 1991; Hansen & Haas, 2001). In an online discussion forum, concurrently posted problems can increase the chances that a potential knowledge provider decides to allocate attention to a focal problem by increasing the salience of the full set of problems and the forum overall. If more problems are posted to the forum on a regular basis, the activity on that forum generally will be greater (see Butler, 2001; Markus, 1987). Potential knowledge providers are more likely to be aware that the forum is an active hub for knowledge sharing, to monitor the forum's postings on an ongoing basis, and to take notice of announcements about new postings. Additionally, they may be more likely to assess the potential for benefits such as reputation enhancement or future reciprocity as greater if there is more activity on the forum (see Chiu et al., 2006; Connolly & Thorn, 1990; Lin, Hung, & Chen, 2009). Whether because a potential knowledge provider is more likely to notice a focal problem or to assess the benefits of responding to it as greater, the result is that more concurrently posted problems can increase the likelihood that the provider allocates attention to the problem.

However, since attention is a finite resource, beyond some point a large number of concurrently posted problems may reduce the likelihood that the potential knowledge provider decides to allocate attention to a focal problem. As March (2002: 27) observed, "the attention devoted to a particular [decision] by a particular potential participant depends on alternative claims on attention." It has long been recognized that problems compete for the attention of members (e.g., Cyert & March, 1963; Simon, 1947), with wide-ranging implications for decision making in organizations (e.g., Eggers & Kaplan, 2009; Joseph & Ocasio, 2012; Sullivan, 2010). The competition between problems has become increasingly acute as companies have introduced new electronic platforms that enable knowledge seekers to "push," or broadcast, their problems to hundreds or thousands of potential knowledge providers at zero marginal cost, simply by posting them to an organization's intranet or external website (Jeppesen & Lakhani, 2010; Shapiro & Varian, 1999). In an online forum, a large number of concurrently posted problems creates many alternative claims on a potential knowledge provider's finite attention. The resulting competitive crowding increases the opportunity costs of attending to a focal problem and thus may decrease the likelihood that the provider decides to allocate attention to that problem.

In summary, when there are few concurrently posted problems, a focal problem can fail to attract attention from a potential knowledge provider owing to a general lack of interest in the forum. Conversely, when there are many concurrently posted problems, a focal problem can fail to attract attention owing to the opportunity costs created by competitive crowding. The implication is that there will be a curvilinear relationship between the number of concurrently posted problems and the allocation of attention to a focal problem, such that a larger number will have positive effects on the likelihood of attention allocation up to a point, after which a larger number will have a negative effect on the likelihood of attention allocation. Thus we predict:

Hypothesis 3. The likelihood that a provider allocates attention to a focal problem will be curvilinearly related to the number of concurrently posted problems, in an inverse U shape.

Moderating Effects of Provider–Problem Expertise Matching

We have argued above that problem length, breadth, and novelty, as well as the number of concurrently posted problems, will influence the likelihood that a potential knowledge provider decides to allocate attention to a focal problem. However, we expect that provider–problem expertise matching will positively moderate these effects. In particular, we argue that closer expertise matching will increase the likelihood that a provider decides to allocate attention to a problem that is longer, broader, or more novel, or which is competing with more concurrently posted problems.

Closer expertise matching can increase the benefits and decrease the costs of allocating attention to a problem of greater length, breadth, or novelty. When a problem is long, broad, or novel, the benefits of allocating attention to that problem will be greater for potential knowledge providers who have expertise that more closely matches the problem, because of their greater ability to offer a response that can help others and create real value, as well as possibly enhance their own reputation and elicit future reciprocity (Chiu et al., 2006; Lin et al., 2009). Additionally, the costs of allocating attention to a problem that is long, broad, or novel are likely to be lower for a potential knowledge provider who has expertise that more closely matches the problem, since that provider will have greater absorptive capacity for the problem, which reduces the costs involved in managing the cognitive load created by length, breadth, or novelty. For example, a provider with more closely matching expertise will be able to sort important from extraneous information in a long problem and digest the important information more efficiently, process the multiple domains of expertise in a broad problem using appropriate schemas more readily, or absorb a problem that is novel for the forum and address that problem more easily.

Closer expertise matching can also increase the benefits and decrease the opportunity costs of allocating attention to a problem when there are more concurrently posted problems competing for the provider's attention. When there are many concurrently posted problems, the benefits of responding to a focal problem will be greater for a provider who has expertise that more closely matches the problem because that provider has greater ability to offer a solution that can create value, be reputation enhancing, and perhaps elicit future reciprocity, compared with a provider who has expertise that is less closely related to the problem. Moreover, the costs of responding to the focal problem will be lower for a provider who has expertise that more closely matches the problem because that provider's time and effort will be more productive as a result of his or her increased capacity to absorb the problem and to articulate a response efficiently (Kotha et al., 2013). Since less time and effort are required to respond to the problem, the opportunity costs incurred by the provider as a result of competitive crowding will be reduced.

Taken together, these arguments suggest that the benefits of allocating attention to a problem that is long, broad, novel, or competing with more concurrently posted problems will be greater when there is a closer expertise match than when there is a distant expertise match, and the costs of allocating attention to that problem will be lower. Accordingly, the inverted U-shaped curves that we predicted for the main effects on attention allocation of problem length, breadth, and novelty, and concurrently posted problems, can be expected to demonstrate a steeper upward curvature and a flatter downward curvature when expertise matching is greater. Hence we predict:

Hypothesis 4. Expertise matching will positively moderate the curvilinear relationship between the likelihood that a provider allocates attention to a focal problem and the problem's (a) length, (b) breadth, and (c) novelty, such that the positive slope of the inverted U-shaped curve becomes steeper and the negative slope becomes flatter with increasing closeness of the provider–problem expertise match.

Hypothesis 5. Expertise matching will positively moderate the curvilinear relationship between the likelihood that a provider allocates attention to a focal problem and the number of concurrently posted problems, such that the positive slope of the inverted U-shaped curve becomes steeper and the negative slope becomes flatter with increasing closeness of the provider–problem expertise match.

DATA AND METHOD

Research Setting

We tested the hypotheses using data collected at one of the world's leading multinational engineering consultancies. Headquartered in London, the firm employs more than 10,000 full-time staff in 71 offices across 26 countries. It executes thousands of projects annually, and is globally renowned for creativity and innovative problem solving through its work on landmark structures including the Sydney Opera House and the 2008 Beijing Olympics National Aquatic Center.

Knowledge sharing enabled engineers in this firm to solve problems that arose from specific client needs and which required them to figure out ways of applying principles, past experience, and existing practices in unique situations. To facilitate knowledge sharing, the firm had invested heavily in advanced information and knowledge management systems, including online discussion forums, as well as expert yellow pages and searchable document repositories. These technology platforms were supplemented by a range of human resources (HR) practices, such as mentoring, job rotation, and experience sharing, as well as by a strong knowledge-sharing culture in which employees were willing to help each other. There were no formal incentives for knowledge sharing, however, and providing advice was not formally rewarded by the appraisal system. Instead, as one senior manager in the firm told us: "People are expected to help . . . so the norm is contribution and this is just the way things are."

Data Collection

To facilitate knowledge sharing across the organization, the firm had established 25 electronic communities of practice (eCOPs), each with its own online discussion forum. These communities focused on different technical disciplines, including structural engineering, fire engineering, environmental consultancy, fluid dynamics, acoustics, etc. Joining an eCOP required individuals to register formally, and registered members received an e-mail whenever a question was posted on the community's online discussion forum. Individuals could belong to multiple eCOPs and could join or leave any of these eCOPs with impunity. The message threads of all of the online discussion forums were accessible by all employees, whether or not they were registered members of the eCOP, and anyone in the firm could contribute to any forum by posting questions and/or answers. The questions and answers posted to a forum included the name of the individuals posting them and their e-mail addresses, but no other identifying information. The system did not allow knowledge providers to automatically access more detailed information on the knowledge seekers. To obtain this information, a provider would have to type the name of the seeker into the search engine of the firm's other knowledge management systems.

We analyzed knowledge sharing in the structural engineering community's online discussion forum over a 32-month period between January 2003 and August 2005. Structural engineering is a fundamental discipline in construction and design projects. The firm employed more than 1,000 structural engineers, who accounted for 27% of its total engineering staff. The structural engineering community was the largest and most vibrant eCOP inside the firm. Like the other 24 eCOPs, it was heavily supported by the organization, which provided funding for video conferences, short courses, lunchtime seminars, and other activities. In August 2005, the structural engineering eCOP had 535 members, of whom 73% were structural engineers, 6% were bridge engineers, 6% were civil engineers, 3% were facade engineers, and others specialized in fields such as material sciences, geotechnical, and infrastructure. The most common themes in the problems that were posted to the structural engineering eCOP's online discussion forum focused on appropriate structural elements, building regulations, economic feasibility, numeric values, and theoretical models and formulae. However, the problems themselves were not titled, tagged, or categorized into these (or other) themes when they were posted. Some sample problems from the online forum are presented in Table 1.

Data sources. We combined data from four sources for this study. The first source was the electronic logs of all of the messages posted to the online discussion forum during the 32-month period under analysis. Because the online discussion forum was used as a vehicle to advertise some of the activities organized by the structural engineering community, such as seminars, workshops, or training courses, we read all of the 3,682 messages posted during our sample period and deleted those messages that did not refer to an engineering problem. After this, the dataset included 3,421 messages, of which 952 were problems and 2,469 were responses. Thus an average of almost 30 problems and 77 responses were posted per month. These messages were posted by 623 individuals, of whom 478 were knowledge providers (that is, posted at least one response).

A second data source was the firm's expert yellow pages. Each member of the firm was encouraged to provide a description of his or her areas of expertise in a personal profile on the company intranet, which could be accessed only by employees, and to keep it updated. These expertise descriptions were self-declared and voluntary. There was a strong incentive to provide an honest and accurate description, because this knowledge management system was searchable and often used by staff to identify experts in a particular area. Indeed, the phrase prompting the expertise description stated "what things I expect people to ring me up and discuss." Thus, while the descriptions were not officially screened for accuracy, an individual was expected to be able to provide an answer to a colleague if questioned about an area of technical expertise listed on his or her profile. Additionally, the descriptions were reviewed annually as part of each individual's appraisal process, which meant that

TABLE 1
Sample Problems Posted on the Structural Engineering Forum

Problem type	Sample post
Appropriate structural elements	"We are undertaking a town center redevelopment and our client is looking to provide the necessary car parking under the development. The site area is approx 30,000m^2 (300m $\times$ 100m). We have a couple of structural options for the deck supporting the development above the car park. A ribbed RC slabs or a steel frame with precast planks. The grid is 16.2 $\times$ 7.4m with an imposed load from the development of 30kN/m^2. Does anyone know of a similar situation and what solution was used for the deck over the car park?"
Building regulations	"We are involved in the design of a football stadium in Scotland. The local building control department has questioned the fact that we haven't got any fire protection to the roof structure. As the roof is not required for the overall stability of the structure, or to hold up any of the floors, we considered that fire protection wasn't required, as in a normal building structure. Has anyone else who has been involved in stadium design had a similar query? Any comments gratefully received."
Economic feasibility	"I am involved in a competition scheme for a housing block right next to a railway, and naturally the architect is concerned about limiting vibration. I know that we have isolated concert halls and the like. However I'm not sure if such measures would be cost-effective in a housing context, and if so, what sort of technologies we might recommend. Any suggestions?"
Numeric values characterizing structural elements	Three questions about shear head reinforcement in flat slabs: 1. "With traditional reinforcement (i.e., straight bars and shear links), what proportion of the reinforcement average weight/square meter would people expect to be accounted for by the shear links?" 2. "7.8m $\times$ 7.8m grid, 300mm flat slab, imposed loads of around 5kPa—what average reinforcement weight per square meter would people expect to see?" 3. "What is the best way of coping with punching shear around columns in flat slabs?"
Theoretical models and formulae	"We are currently designing a number of high rise apartment blocks in masonry which exceed 4 stories. Walls are load bearing masonry with precast floors. We are currently designing the buildings to Option 3 of Table 12 BS 5628 i.e., designed vertical and horizontal ties for accidental damage. This is the Client's preferred option. The horizontal ties are not a problem. His preferred method of forming the vertical ties is to use a hollow block which is then in filled with concrete. When you use the formula in Table 14 BS 5628 to calculate the tie force—for a 150mm thick inner leaf with ties at 5m centers and a clear distance between floor restraints of 2.6m—it works out at approximately 1MN. This equates to approximately 5T25's. I have looked through the Masonry Designers Manual which comes up with 4T32's in their example. The values appear high. If anybody has used this method before and can provide any advice on the above, I would be grateful."

there was some formal, as well as informal, pressure on staff not to "overdeclare" their expertise. About two-thirds of the firm's employees had completed their expertise descriptions when we obtained this dataset. These descriptions were 30 words long on average, although some exceeded 250 words. They provided rich information, as this typical example demonstrates:

> Structural issues related to reinforced/pre-stressed and/or post tensioned concrete; flat slab/rib slab design; in service behaviour including deflection prediction, structural implications of shrinkage and thermal effects, and the investigation of defect; 3D steelwork package Xsteel and Raft design; structural testing and monitoring including full scale testing of hole cutting in a post-tensioned slab [56 words]

The third data source was records from the HR department, which provided data on each employee's office location, rank, tenure, and gender. The fourth source was the company's project database, from which we extracted lists of the projects on which each individual had worked since joining the firm.

Because of missing data across the four data sources, some of the knowledge providers had to be dropped from the dataset, reducing our final sample to 307 knowledge providers (although our models also account for individuals who could have served as knowledge providers, but did not, as described below). When we compared these 307 providers with the 171 providers who were excluded because of missing data, we found that those who were included in our final sample posted significantly more answers than those who were excluded ($p < .01$). This is advantageous for our study in that we are able to include a high proportion of responses that were posted to the forum in the final sample even though some individuals had to be dropped. In fact, the sample of 307 knowledge providers was responsible for generating 76% of all of the responses to engineering-related problems on the forum during the period under analysis. After excluding responses to problems posted by individuals with missing data, our final sample included 1,974 messages, of which 639 were problems and 1,336 were responses.[2]

Statistical Approach

Because the focus of our theoretical arguments is on whether a particular knowledge provider decides to allocate attention to a particular problem in the forum, the unit of analysis in our main econometric models is a provider–problem dyad. We constructed a matrix of all provider–problem dyads in which the *ij*th cell is 1 if provider *i* provided a response to problem *j* (realized dyad) or 0 if provider *i* did not provide an answer to problem *j* (nonrealized dyad). The providers in these dyads included all of the 307 individuals who posted at least one response during the observation period. We defined the risk set of problems to include all possible problems that were available to be answered at the time that a focal problem was posted to the forum—that is, as all problems posted prior to the time when the focal problem was posted (whether or not they received a response) and which were still open at the time when the focal problem was posted. We considered a problem to be still open if it was posted less than 50 days before the focal problem; we used this window because no problem in our dataset received a response 50 days or more after it been posted. The resulting dataset consisted of 376,670 possible provider–problem dyads, of which 1,336 were coded 1 (realized dyads) and 375,334 were coded 0 (nonrealized dyads).

Constructing the dataset in this way enabled us to compare realized dyads with nonrealized dyads, following the analytic approach taken in previous studies of tie formation between firms (e.g., Gulati, 1995). However, the dataset was characterized by a preponderance of zeros resulting from the large number of nonrealized dyads. The analysis of a dataset with very few positive events (less than 1%) cannot be undertaken using a standard logit model because it will underestimate the probability of a positive outcome—that is, a match between a provider and a problem (King & Zeng, 2001). The dataset was also characterized by nonindependence in the error terms arising from the fact that both providers and problems could appear many times in the dataset. This issue of network autocorrelation could lead to underestimation of standard errors (Krackhardt, 1988). To address these concerns, we followed previous studies of tie formation in sparse networks (e.g., Hallen, 2008; Jensen, 2003) by using a choice-based sampling technique and testing our hypotheses using a rare-event logit model. The choice-based sampling technique included all of the realized dyads and a randomly extracted sample of corresponding nonrealized dyads.

Consistent with our theoretical focus on why a particular provider decides to allocate attention to a particular problem rather than to other possible problems, for each realized dyad in which provider *i* responded to a problem *j*, we randomly selected 10 nonrealized dyads from the sample of problems to which provider *i* could have responded, but did not (that is, those posted less than 50 days prior to the focal problem). To ensure that enough problems had been posted prior to the focal problem to randomly extract the sample of 10 nonrealized dyads, we excluded problems posted during the first two months of our observation period, resulting in a final dataset with 13,761 dyads, of which 1,251 were realized dyads and 12,510 were nonrealized dyads.[3]

While this choice-based sampling technique resolves concerns created by a preponderance of zeros in the dataset, it can bias the logit estimates because

[2] It is possible that some responses to problems were given directly to a seeker, bypassing the online forum. However, our interviews with members of the firm, including the head of knowledge management, indicated that participants were strongly encouraged to post their responses on the forum rather than to reply directly to a seeker, so that others could search the forum for answers to their questions. Indeed, seekers occasionally posted on the forum answers that they had received over the phone from a provider for exactly that purpose. We also checked whether the seeker had included a phone number or e-mail address at the end of the question, which could indicate that a direct, rather than public, response was desired, and found only two instances of this in our sample of 639 problems. The available evidence thus indicates that providers tended to post their responses on the forum rather than to reply directly to seekers via e-mail or telephone.

[3] We ran robustness tests with ratios of 1:5 and 1:3 realized to nonrealized dyads, and found that they produced substantively equivalent results to those reported here.

the proportion of positive outcomes in the sample is different from that in the underlying population of potential dyads. To correct this bias, we used weighted exogenous sampling maximum-likelihood estimation (WESMLE), an approach that weights the contribution of each dyad to the likelihood function and is better than alternative approaches for large samples (King & Zeng, 2001). Additionally, we clustered the standard errors on the provider (Hallen, 2008; Jensen, 2003), since each provider appears in one realized dyad and 10 nonrealized dyads (that is, the provider is constant across 11 observations).[4] We used the ReLogit Stata procedure of Tomz (2003) to estimate the logit models. Finally, we utilized the longitudinal nature of the dataset by constructing the explanatory and control variables to minimize reverse causality by measuring them in the period prior to the focal match/nonmatch, as explained more fully below.

Dependent Variable

Attention allocation. Our main dependent variable is whether a provider decided to allocate attention to a problem posted on the structural engineering community's online discussion forum. We considered that provider *i* allocated attention to problem *j* if he or she posted a response to problem *j*. Thus our dependent variable (*attention allocation*) is a binary variable that is equal to 1 if a possible provider–problem match was realized, or 0 if that possible match was not realized.

Explanatory Variables

Provider–problem expertise matching. To capture how close the expertise possessed by a provider was to the expertise required by the problem, we utilized a keyword similarity approach (Criscuolo, Salter, & Sheehan, 2007). Specifically, our measure of *expertise matching* was constructed by capturing how similar the keywords in the provider's expertise description were to the keywords in the focal problem.

To capture the universe of possible keywords and how similar they were to each other, we began by deriving a list of 574 keywords from the 3,948 expertise descriptions of all employees stored in the company's expert yellow pages.[5] We used this list of 574 keywords to construct a keyword-by-keyword similarity matrix (K) (574 $\times$ 574), the *ij*th cell of which contains a measure of similarity between keywords *i* and *j*. To derive this measure of similarity, we used the Salton cosine formula:

$$\text{cosine(i,j)} = \frac{\text{cooc(i.j)}}{\sqrt{\text{oc(i)*oc(j)}}}$$

where the nominator represents the co-occurrence of each pair of keywords in the expertise descriptions and the denominator is the product of the square root of the respective occurrence frequencies in all 3,948 expertise descriptions (see Aral & Van Alstyne, 2011, for a similar application in the context of e-mail exchanges). Pairs of keywords that coappear very often have a cosine nearer to 1, while keywords that rarely appear together have a cosine nearer to 0. For example, the cosine value for "foundation" and "pile" is 0.46, while the cosine between "foundation" and "vibration" is only 0.027.

We also used the list of 574 keywords to construct a provider-by-keyword asymmetric matrix (X) (307 $\times$ 574), in which cell $x_{ij} = 1$ if the *i*th provider mentioned keyword *j* in his or her expertise description, and $x_{ij} = 0$ otherwise. Similarly, we constructed a problem-by-keyword asymmetric matrix (Y) (639 $\times$ 574), in which cell $y_{ij} = 1$ if problem *i* mentioned keyword *j*, and $y_{ij} = 0$ otherwise. We then multiplied the provider-by-keyword matrix (X) by the keyword similarity matrix (K), and multiplied the resulting matrix by the transposed problem-by-keyword matrix (Y). By weighting by the keyword similarity matrix, we were able to capture the extent of similarity between the keywords in the provider's

[4] We also estimated models clustering the errors on the problem, the knowledge seeker, and both the knowledge provider and the knowledge seeker, and found results consistent with those reported here.

[5] In deriving this list of keywords, we disregarded articles, prepositions, adverbs, verbs, and words that did not refer to technical expertise. We also classified word pairs, such as "remote sensing" and "traffic calming," and word triplets such as "environmental impact assessment" and "computational fluid dynamics," as keywords. Additionally, keywords were corrected for plurals and association, e.g., "rail/railway," "sustainable/sustainability," "daylight/light," "cabling/cable," and "forecasting/forecast." (For an application of this approach to the context of patent analysis, see Corrocher, Malerba, & Montobbio, 2007.) From this list, we selected the 574 keywords that appeared more than 10 times. We then presented this list to senior managers to ensure that key areas of expertise were not missing, and that the list of pairs and triplets of keywords did identify particular areas of expertise.

expertise and the keywords in the problem, even when these keywords were not exactly the same. In this way, we obtained a provider-by-problem matrix (W) (307 × 639), which contained in cell w_{ij} the similarity between the keywords mentioned in the expertise description of provider *i* and those mentioned in problem *j*. We then divided the value of each w_{ij} cell by the product of the total number of keywords in the expertise description of provider *i* and the total number of keywords in problem *j*, to restrict the range of this indicator between 0 and 1. Problems that addressed areas of expertise more similar to the expertise of the potential knowledge provider have a higher value of this expertise matching variable.

Problem characteristics. To measure *problem length*, we counted the number of words in each problem posted to the forum.

To measure *problem breadth*, we computed the extent to which there was variety in the domains of expertise addressed in the problem. To identify the possible domains of expertise that could be addressed, we again drew on the 574 keywords from the company's expert yellow pages. We carried out a hierarchical clustering analysis on the keyword-by-keyword matrix (K), applying the Ward method with Euclidean distances. Using the stopping rule of Duda and Hart (1973), we obtained 19 clusters of keywords, which represented different domains of expertise inside the company. This method allowed us to classify the keywords that appeared in each problem into one or more of these 19 domains of expertise. We constructed the measure of problem breadth using Teachman's entropy index, a measure of variety (see Harrison & Klein, 2007) determined by the following formula:

$$\text{problem breadth}_i = \sum_{i=1}^{19} p_i \times \ln(p_i)$$

where p_i is the proportion of keywords in domain *i*. Problems in which keywords are spread more evenly across a higher number of expertise domains have a higher value on this breadth measure.

Our measure of *problem novelty* captures how different the focal problem is from problems previously posted to the forum, again using a keyword similarity approach to derive the measure. Specifically, the measure was constructed by examining how similar the keywords in the focal problem were to the keywords in previously posted problems. To capture how similar a particular problem was to each previously posted problem, we multiplied the problem-by-keyword matrix (Y) by the keyword similarity matrix (K), and then multiplied the resulting matrix by the transposed problem-by-keyword matrix (Y) to generate a problem-by-problem matrix (Q), which contained in cell q_{ij} the similarity between the keywords in problem *i* and those in problem *j*. We then divided the value of each q_{ij} cell by the product of the total number of keywords in problem *i* and in problem *j*, to account for all of the possible combinations of keywords in two given problems. Finally, for each problem *i* we calculated the average similarity value between problem *i* and all of the other problems previously posted on the forum (that is, we excluded problems posted after the focal problem), and computed the inverse of this average to derive our problem novelty measure.[6] Accordingly, for a given problem *i* the problem novelty variable is derived using the following formula:

$$\text{problem novelty}_i = \frac{1}{\frac{\sum_{j=1}^{J} \frac{q_{ij}}{kw_i^* kw_j}}{j}}$$

where *J* is the number of problems previously posted on the forum.

Thus a problem that contains keyword combinations that differ from the keyword combinations in previously posted problems will score highly on this novelty measure. By construction, problems posted at the beginning of our observation period will tend to display lower values on this measure than problems posted towards the end of the period. To address this issue, we included month and year dummies in our models that account for the timing of the problems.[7]

Problem crowding. We constructed a measure of *concurrent problems* that is equal to the number of problems posted on the forum in the three working days prior to the focal problem being posted. We chose a window of three working days because close to 90% of the problems in our dataset were answered within this time frame. We ran robustness checks with different windows, including five, seven, and 10 working days, and obtained similar results.

[6] Consistent results were also obtained with a problem novelty measure built using all 639 problems posted to the forum during our observation period, including those posted before, as well as after, the focal problem.

[7] We also reran our analyses after dropping the first six months of observations, and the results did not change.

Control Variables

We included several sets of variables to control for alternative explanations for attention allocation in the online discussion forum. First, we included a series of variables to account for characteristics of the provider–seeker dyad. To control for *reciprocity* (that is, the possibility that a provider might be more likely to respond to a problem posted by a seeker who had previously assisted him or her), we used a dummy variable equal to 1 if the focal provider had previously received a response to a problem from the focal seeker, or 0 otherwise. To control for homophily (that is, the possibility that a provider might be more likely to respond to a problem posted by a seeker who shares similar personal characteristics), we created a dummy variable (*same gender*) equal to 1 if both individuals in a dyad were of the same gender, or 0 otherwise.[8] To control for proximity (that is, the possibility that a provider might be more likely to respond to a problem posted by a seeker in the same location), we included a dummy variable (*shared office*) equal to 1 if two individuals worked in the same office, or 0 otherwise. To control for familiarity (that is, the possibility that a provider might be more likely to respond to a problem posted by a seeker whom he or she knew), we included two variables: a dummy variable (*shared projects*) equal to 1 if two individuals had worked together on a project during the five years preceding the date on which the problem was posted on the forum, or 0 otherwise; and a count variable (*shared communities*) that captures the number of other online communities in which the provider and seeker were both members, since this could have enabled them to get to know each other through interactions on other online discussion forums, as well as through other community-related activities, such as video conferences, seminars, and training sessions.

Second, we included a series of variables to account for characteristics of the seeker that might lead a provider to allocate attention to a problem posted by that seeker. We controlled for the *seeker's rank* in the company using HR data that classified each individual's hierarchical level on a nine-point scale (1 = junior consultant, 9 = director), to capture the possibility that a provider might be more likely to allocate attention to a problem posted by a seeker with higher rank in the company. We included a dummy variable (*seeker member*) equal to 1 if the seeker was a member of the structural engineering community, or 0 otherwise, since a provider might have felt more motivated to respond to a problem posted by a seeker who was more invested in the forum. We also included a dummy variable (*seeker facilitator*) equal to 1 if the seeker was one of the formal facilitators in the online discussion forum, or 0 otherwise. These formal facilitators were subject matter experts responsible for stimulating technical discussions and maintaining an active discussion forum; given their central role in the community, problems posted by them might have been more likely to attract the attention of a knowledge provider.

Third, we accounted for characteristics of the provider that might have influenced his or her decision to allocate attention to a particular problem. We controlled for the *provider's rank* and the *provider's tenure* in the organization. Individuals in higher positions in the company and/or with longer tenure might have had a greater depth of expertise in particular areas, which could have increased their propensity to respond to problems in those areas. We also controlled for the number of projects (logged) to which a provider was assigned at the time that the problem was posted to the forum (*provider project load*), because a provider who was working on more projects at the time a problem was posted to the forum might have been less likely to allocate attention to that problem as a result of his or her higher project load.

Fourth, to account for the possibility that a particular provider was not the first to respond to a problem, we included a control variable (*response order*) for the order of his or her response (that is, first, second, third, etc.). We expected a provider to be less likely to respond to a problem if others had already responded. We constructed this variable by setting its value equal to the actual order of the response for a problem to which a provider responded; for a problem to which a provider could have responded, but did not, we randomly assigned a value to this variable, so that its distribution among the nonrealized dyads corresponded to that among the realized dyads.

[8] We would have collected information on other demographic characteristics, such as age, race, or nationality, if it had been possible, but we were constrained by laws that restricted the use of such information. However, while we intuitively expected that gender might matter in an online discussion forum, because it is often apparent from participants' names, these other demographic characteristics are less likely to matter significantly; in support of this assumption, prior research has shown limited roles for their effects in online knowledge sharing (Constant et al., 1996).

Fifth, we included dummy variables for years, months, and days of the week in our models to account for any otherwise unobserved tendency of knowledge providers to allocate attention to problems posted at different times.

Controlling for Selection Bias

Consistent with the focus of our theoretical arguments on why a particular provider chooses to allocate attention to a particular problem rather than to other possible problems in the online discussion forum, our main rare-event logit analyses focus on the 307 individuals who posted at least one response to a problem on the forum during our observation period. However, restricting our analysis to only those individuals who acted as knowledge providers creates a possible selection bias, because individuals who post responses to problems may systematically differ from individuals who do not post responses. To account for this possible bias, we used the two-stage procedure proposed by Heckman (1976). In our context, this involved estimating a first-stage probit model to predict whether an individual posted at least one response to any problem on the forum during the observation period (the "selection model"). From this, we derived an inverse Mills' ratio, which we then included in our main rare-event logit model (the "outcome model").

The sample used in the selection model included all 399 individuals who were active on the forum during the observation period, whether as knowledge providers, knowledge seekers, or both, plus all 214 members of the structural engineering community who were not active on the forum during the observation period and for whom we had complete data, for a total risk set of 613 individuals.[9] The dependent variable (*knowledge provider*) was equal to 1 if an individual posted a response on the forum at any point during the observation period, or 0 otherwise. As independent variables, we included a series of characteristics that we expected might have influenced whether an individual was a knowledge provider. Individuals who were formal *forum facilitators* of the forum may have been more likely to respond to problems posted on the forum as part of their responsibilities. Individuals who were *members of the structural engineering (focal) community* may have had a greater sense of commitment to the community and thus have been more likely to respond to problems, while individuals who were *members of (more) other communities* may have had a greater underlying propensity to share knowledge and to help others by posting responses. *Gender* might have affected the probability of knowledge provision too, so we included a dummy variable that was equal to 1 for male and 0 for female. Individuals in higher positions in the company (*rank*) were often expert problem solvers and may have been more able to respond to problems on the forum as a result, while individuals who had worked in the company for a longer period (*tenure*) had more work experience that could be shared with others on the forum. Individuals who had expertise in a larger number of engineering domains (*expertise breadth*) may have been more likely to have the relevant knowledge to respond to problems on the forum; we calculated this variable following the procedure used to derive the problem breadth variable. Similarly, individuals who specialized in structural engineering (*structural engineer*) may have had a greater propensity to respond to problems on the forum. Finally, individuals who were assigned to more projects during the observation period may have been less likely to respond to problems posted on the forum as a result of their higher project load, so we included a logged measure of total project assignments in the model (*total project load*).[10]

[9] We also ran the first-stage model using as a risk set all individuals who were active in the forum during our observation period *plus* all structural engineers in the company. This risk set includes 955 individuals, of whom 543 did not participate in the forum. For this larger risk set, we do not have information on gender and project load for all individuals. Since both of these variables are significant in the current model, we have chosen to estimate our first-stage model using the smaller sample. The results do not change with the larger risk set.

[10] To apply the Heckman two-stage procedure, we need at least one instrumental variable that is expected to influence the selection process, i.e., whether an individual acts as knowledge provider, but is not expected to influence the likelihood that an individual allocates attention to a particular problem rather than to other problems. We used the following four variables as instruments: *forum facilitator*, *member of focal community*, *member of other communities*, and *gender*. We included *rank* and *tenure* in both the first-stage model and our main models, since these variables could be expected to influence both the selection and the outcome equations. We replaced *expertise breadth* and *structural engineer* with our expertise-matching measure in the main models, since this measure more accurately captures how expertise might affect a provider's decision to allocate attention to a particular problem (the results are not changed by including them too). We also replaced *total project load* with the measure that captures the provider's project load at the time that the focal problem was posted.

RESULTS

Table 2 reports the descriptive statistics for the variables used in the models. We standardized the main continuous independent variables by subtracting the mean and dividing by the standard deviation, in order to avoid high correlations between these variables and their interaction terms (Neter, Wasserman, & Kutner, 1990). Most of the correlation coefficients are low. Nevertheless, we derived variance inflated factors (VIF) for our models; these were on average less than 2.5, indicating that multicollinearity is not a concern in the regressions.

The first-stage probit model predicting whether or not an individual acted as a knowledge provider (not shown) indicated that individuals were more likely to post at least one response on the online discussion forum during the observation period if they were higher ranked in the company (β = .11, $p < .01$), were male (β = .30, $p < .01$), had broader expertise (β = .31, $p < .01$), were structural engineers (β = .21, $p < .05$), and—contrary to our expectations—had a higher project load (β = .20, $p < .05$). The other variables included in the model did not have significant effects.

The results of our main rare-event logit models are shown in Tables 3 and 4. Table 3 reports the estimates for the main effects predicted by Hypotheses 1, 2, and 3. Model 1 is a baseline model that includes only the control variables. This model indicates that reciprocity was a positive and significant predictor of the probability that a particular provider allocated attention to a particular problem. However, we did not find any significant effects for homophily based on gender, proximity based on shared office, or familiarity based on shared projects or shared communities. Similarly, none of the seeker characteristics seemed to explain why a provider decided to allocate attention to a problem. Of the provider characteristics, rank had a positive and significant effect in the model with control variables only, but this effect is not significant in subsequent models. Conversely, project load is not significant in the model with control variables only, but this variable is negative and significant in subsequent models, indicating that providers were less likely to allocate attention to a focal problem if they were assigned to more projects at the time that the problem was posted. The inverse Mills' ratio is not significant, indicating that selection bias was not a major concern in our dataset.

Model 2 shows the results for Hypothesis 1, which predicted that a provider will be more likely to respond to a problem that more closely matches his or her expertise. This hypothesis is supported: we find a positive and significant relationship between expertise matching and the decision to allocate attention to a given problem (β = .17, $p < .01$). Models 3, 4, and 5 add the problem length, breadth, and novelty variables in order to test Hypotheses 2a, 2b, and 2c, which predicted curvilinear relationships between each of these three problem characteristics and the likelihood that a provider decides to allocate attention to the problem. The predictions are supported for all three variables, as shown by the signs and significance of the coefficient estimates. In Model 3, we find a positive and significant coefficient for the problem length variable (β = .15, $p < .01$), and a negative and significant coefficient for the squared term (β = –.08, $p < .01$). In Model 4, the coefficient for the problem breadth variable is also positive and significant (β = .64, $p < .01$), and its squared term is negative and significant (β = –.31, $p < .01$). The same pattern is found in Model 4 for the problem novelty variable, which has a positive and significant linear effect (β = .10, $p < .05$), and a negative and significant squared term (β = –.03, $p < .05$). The inflection points for all three inverted U-shaped curves are within the observed range of these variables. These results hold in Model 6, in which the variables are included together. Hence we conclude that there is strong support for Hypothesis 2.

In Model 7, we introduce the linear and squared terms for the competing problems variable in order to test Hypothesis 3, which predicted that the number of concurrently posted problems has a curvilinear relationship with the likelihood that a provider decides to allocate attention to a focal problem. In partial support of Hypothesis 3, we find that the linear term has the predicted positive and significant effect on the likelihood of attention allocation (β = .11, $p < .01$). However, the squared term is not negative and significant, as we had predicted; instead, it is positive and nonsignificant. The finding of a positive linear effect of the number of competing problems holds in Model 8, in which the nonsignificant squared term is excluded, indicating that the likelihood that a provider allocated attention to a focal problem was greater if a higher number of other problems were posted to the forum concurrently, and did not decline as the number of concurrently posted problems reached higher levels. Thus the support for Hypothesis 3 is mixed.

To test Hypotheses 4 and 5, the moderating effects of expertise matching on problem characteristics

TABLE 2
Descriptive Statistics and Bivariate Correlations

	Variable	Mean	*SD*	Min.	Max.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	Attention allocation	0.09	0.29	0	1																	
2	Reciprocity	0.00	0.04	0	1	0.05																
3	Same gender	2.82	2.11	1	15	−0.01	0.00															
4	Shared office	5.68	1.99	1	9	0.01	0.02	−0.01														
5	Shared communities	0.90	0.30	0	1	0.00	0.01	0.00	−0.06													
6	Shared projects	0.06	0.25	0	1	0.02	0.00	0.01	0.37	0.09												
7	Seeker rank	0.82	0.39	0	1	0.00	0.02	0.00	0.06	−0.00	0.07											
8	Seeker member	0.19	0.39	0	1	−0.00	−0.00	0.01	0.02	−0.00	−0.07	−0.10										
9	Seeker facilitator	0.19	0.60	0	5	0.01	−0.01	0.00	−0.03	0.09	−0.04	0.01	0.09									
10	Provider rank	0.34	0.47	0	1	0.01	0.00	0.01	0.14	0.02	0.08	−0.02	0.37	0.13								
11	Provider tenure	6.50	1.82	2	9	0.00	0.01	−0.01	0.03	0.00	0.01	0.01	0.09	0.06	0.23							
12	Provider project load	18.61	11.08	3	53	0.00	−0.00	−0.00	0.02	0.00	0.02	0.01	0.13	0.07	0.34	0.55						
13	Response order	18.98	20.10	0	93	−0.00	0.03	0.00	0.02	−0.01	0.01	0.03	0.15	0.17	0.43	0.40	0.66					
14	Expertise match	0.05	0.04	0	1	0.10	0.01	0.01	−0.02	−0.00	0.01	0.03	0.00	0.00	0.06	−0.01	0.02	0.10				
15	Problem length	97.48	66.17	9	540	0.00	0.00	−0.00	−0.21	0.04	−0.14	0.07	−0.01	0.05	−0.05	0.01	0.01	−0.01	−0.00			
16	Problem breadth	0.98	0.55	0	2.03	0.01	0.00	0.00	−0.11	0.02	−0.08	0.08	0.00	0.01	−0.04	0.01	0.01	0.00	0.01	0.45		
17	Problem novelty	17.88	8.99	0	103.62	−0.01	−0.00	−0.00	0.13	−0.06	0.05	−0.01	0.03	0.01	0.03	0.01	0.00	−0.01	−0.27	−0.05	−0.05	
18	Competing problems	3.95	2.11	0	12	0.02	0.01	0.01	0.00	−0.00	−0.08	−0.03	0.04	0.01	0.02	0.00	−0.00	0.01	−0.02	0.01	−0.02	−0.02

Note: Correlations greater than 0.02 are significant at 5%; n = 13,761.

TABLE 3
Rare Event Logit Model Estimations for Hypotheses 1, 2, and 3

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Reciprocity	2.29***	2.28***	2.23***	2.29***	2.29***	2.25***	2.27***	2.25***
	(0.51)	(0.52)	(0.54)	(0.55)	(0.53)	(0.56)	(0.57)	(0.57)
Same gender	0.04	0.01	−0.00	−0.00	0.01	0.01	0.01	0.01
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Shared office	−0.06	−0.05	−0.05	−0.05	−0.05	−0.05	−0.05	−0.05
	(0.08)	(0.09)	(0.08)	(0.09)	(0.09)	(0.09)	(0.08)	(0.08)
Shared communities	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Shared projects	0.07	0.06	0.07	0.06	0.06	0.07	0.07	0.07
	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Seeker rank	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Seeker member	0.06	0.05	0.05	0.05	0.05	0.06	0.06	0.06
	(0.10)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
Seeker facilitator	0.06	0.06	0.11	0.09	0.05	0.11	0.14	0.15
	(0.13)	(0.14)	(0.13)	(0.14)	(0.14)	(0.13)	(0.14)	(0.13)
Provider rank	0.02**	0.02	0.01	0.01	0.02	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Provider tenure	−0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Provider project load	−0.01	−0.04**	−0.04**	−0.04**	−0.04**	−0.04**	−0.04**	−0.04**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Response order	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Expertise matching[a]		0.17***	0.17***	0.17***	0.17***	0.18***	0.18***	0.18***
		(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
Problem length[a]			0.15***			0.14***	0.14***	0.14***
			(0.05)			(0.05)	(0.05)	(0.05)
Problem length^2			−0.08***			−0.07***	−0.07***	−0.07***
			(0.03)			(0.03)	(0.03)	(0.03)
Problem breadth[a]				0.64***		0.47***	0.46**	0.44**
				(0.19)		(0.18)	(0.18)	(0.18)
Problem breadth^2				−0.31***		−0.27***	−0.26**	−0.25**
				(0.11)		(0.10)	(0.10)	(0.10)
Problem novelty[a]					0.10**	0.10*	0.09*	0.09*
					(0.05)	(0.05)	(0.05)	(0.05)
Problem novelty^2					−0.03**	−0.02**	−0.02**	−0.02**
					(0.01)	(0.01)	(0.01)	(0.01)
Competing problems[a]							0.11***	0.13***
							(0.04)	(0.03)
Competing problems^2							0.02	
							(0.02)	
Inverse Mills' ratio	0.09	−0.02	−0.03	−0.03	−0.01	−0.03	−0.02	−0.02
	(0.06)	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)	(0.09)	(0.09)
Constant	−5.30***	−5.15***	−5.08***	−5.38***	−5.05***	−5.18***	−5.04***	−4.93***
	(0.49)	(0.51)	(0.51)	(0.51)	(0.50)	(0.51)	(0.54)	(0.51)
Log-likelihood	−4136.35	−4090.10	−4083.70	−4084.90	−4084.65	−4076.15	−4068.70	−4069.10

Note: Robust standard errors clustered by providers in parentheses; year, month, and day of the week dummies included; DV = attention allocation to a focal problem; n = 13,761.

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

* Significant at $p < .10$

** Significant at $p < .05$

*** Significant at $p < .01$

TABLE 4
Rare Event Logit Model Estimations for Hypotheses 4 and 5

	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18
Reciprocity	2.21***	2.22***	2.19***	2.19***	2.21***	2.22***	2.25***	2.23***	2.20***	2.16***
	(0.57)	(0.57)	(0.57)	(0.57)	(0.57)	(0.57)	(0.57)	(0.57)	(0.56)	(0.56)
Same gender	0.00	0.00	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Shared office	−0.05	−0.05	−0.05	−0.05	−0.05	−0.05	−0.05	−0.05	−0.04	−0.04
	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	(0.08)
Shared communities	0.06	0.06	0.06	0.06	0.05	0.05	0.06	0.06	0.06	0.05
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Shared projects	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.04
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Seeker rank	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Seeker member	0.06	0.06	0.06	0.06	0.08	0.08	0.08	0.08	0.08	0.08
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
Seeker facilitator	0.12	0.12	0.15	0.15	0.15	0.15	0.09	0.11	0.12	0.13
	(0.14)	(0.14)	(0.13)	(0.13)	(0.13)	(0.13)	(0.14)	(0.14)	(0.14)	(0.13)
Provider rank	0.02	0.02	0.02	0.02	0.01	0.01	0.02	0.02	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Provider tenure	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Provider project load	−0.05**	−0.05**	−0.06**	−0.06**	−0.04*	−0.04*	−0.04*	−0.04*	−0.05**	−0.05*
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Response order	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Expertise matching (EM)[a]	0.28***	0.28***	0.12***	0.13***	0.30***	0.28***	0.25***	0.25***	0.21***	0.27***
	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.07)	(0.05)
Problem length[a]	0.09	0.09	0.15***	0.15***	0.14***	0.14***	0.14***	0.14***	0.13**	0.12**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Problem length ^2	−0.06**	−0.06**	−0.07***	−0.07***	−0.07***	−0.07***	−0.07***	−0.07**	−0.07***	−0.07**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Problem breadth[a]	0.40**	0.40**	0.23	0.28	0.42**	0.43**	0.42**	0.40**	0.29	0.30*
	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)
Problem breadth^2	−0.23**	−0.23**	−0.17*	−0.20**	−0.25**	−0.25**	−0.24**	−0.23**	−0.20*	−0.20*
	(0.11)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)	(0.10)	(0.10)	(0.10)
Problem novelty[a]	0.13**	0.13**	0.14***	0.14***	0.12**	0.12**	0.12**	0.13**	0.15***	0.15***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Problem novelty^2	−0.03**	−0.03**	−0.03***	−0.03***	−0.02	−0.01	−0.02**	−0.02**	−0.01	−0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Competing problems[a]	0.13***	0.13***	0.12***	0.12***	0.13***	0.13***	0.08**	0.10***	0.09**	0.11***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
Competing problems^2							0.03		0.02	
							(0.02)		(0.03)	

TABLE 4
(Continued)

	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18
EM × Problem length	0.14***	0.14***							0.05	0.07*
	(0.04)	(0.04)							(0.04)	(0.04)
EM × Problem length ^2	−0.00								0.03	
	(0.02)								(0.02)	
EM × Problem breadth			0.32***	0.20***					0.16	0.10**
			(0.12)	(0.04)					(0.13)	(0.05)
EM × Problem breadth ^2			−0.09						−0.03	
			(0.09)						(0.09)	
EM × Problem novelty					0.13***	0.13***			0.10***	0.10***
					(0.03)	(0.03)			(0.03)	(0.03)
EM × Problem novelty ^2					−0.01				0.01	
					(0.02)				(0.02)	
EM × Competing problems							0.09***	0.09***	0.05**	0.06**
							(0.02)	(0.02)	(0.02)	(0.02)
EM × Competing problems ^2 squared							−0.01		0.01	
							(0.01)		(0.02)	
Inverse Mills' ratio	0.01	0.01	−0.01	−0.02	0.00	0.00	0.03	0.03	0.02	0.02
	(0.12)	(0.12)	(0.13)	(0.13)	(0.10)	(0.10)	(0.11)	(0.11)	(0.13)	(0.13)
Constant	−5.01***	−5.01***	−4.88***	−4.90***	−4.96***	−4.96***	−5.21***	−5.04***	−5.09***	−4.97***
	(0.53)	(0.53)	(0.54)	(0.53)	(0.52)	(0.52)	(0.55)	(0.52)	(0.57)	(0.53)
Log-likelihood	−4063.28	−4063.49	−4057.76	−4058.11	−4057.79	−4058.32	−4065.18	−4065.66	−4046.42	−4047.44

Note: Robust standard errors clustered by providers in parentheses; year, month, and day of the week dummies included; DV = attention allocation to a focal problem; $n = 13{,}761$.

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

* Significant at $p < .10$

** Significant at $p < .05$

*** Significant at $p < .01$

and problem crowding are presented in Table 4, in Models 9–14 and 15–16, respectively. As these models show, we find linear-by-linear interactions, but not curvilinear-by-linear interactions, between expertise matching and each of the variables capturing problem characteristics and problem crowding—that is, there are significant interactions with the linear terms, but not the squared terms, for these variables. Model 17 presents a full model that includes the interactions for the squared terms, while Model 18 reports the full model in which we excluded these higher-order interaction terms. Because the interactions for the squared terms are not significant in the partial models, we assess the support for our moderating hypotheses using Model 18 (see Aiken & West, 1991).

Model 18 shows a positive and significant interaction term between expertise matching and problem length ($\beta = .07$, $p < .10$): the likelihood that a provider allocated attention to a longer problem was greater if that provider had expertise that more closely fit the expertise required by that problem. The interaction between expertise matching and problem breadth is also positive and significant ($\beta = .10$, $p < .05$), indicating that the likelihood that a provider allocated attention to a broader problem was higher if that provider had expertise that more closely matched that required by the problem. Similarly, expertise matching positively and significantly moderates the relationship between problem novelty and attention allocation ($\beta = .10$, $p < .01$). We also find that the interaction term between expertise matching and competing problems ($\beta = .06$, $p < .05$) is positive and significant, indicating that the likelihood that a provider allocated attention to a focal problem while facing a higher number of concurrently posted problems was greater if that provider's expertise matched the expertise called for by the problem.

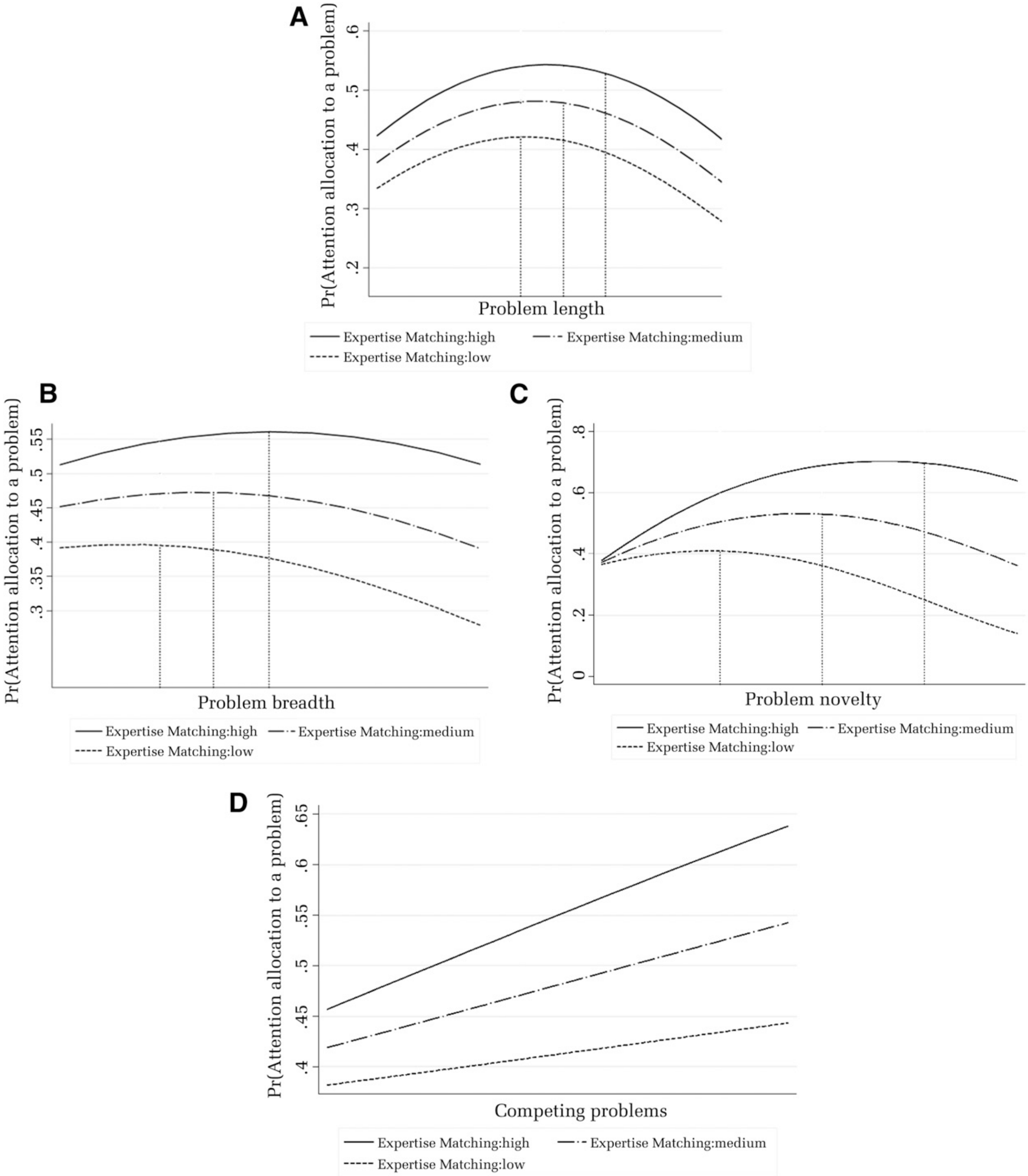
Notably, although there are no significant interactions for the squared terms, it is still possible that the negative slopes of the curvilinear main effects may become flatter with increasing closeness of the provider–problem expertise match. To see this, consider the derivative for a linear-by-linear interaction in a simple linear model, $Y = \beta_1 X + \beta_2 X^2 + \beta_3 Z + \beta_4 XZ$, where X is problem length and Z is expertise matching. The derivative, $dY/dX = \beta_1 + 2\beta_2 X + \beta_4 Z$, shows that the slope of the curve is a function of both X and Z—that is, both the upward-sloping and downward-sloping parts of the curve are affected by Z (see Aiken & West, 1991, for further explication).[11] In order to establish whether expertise matching significantly affects both the positive and the negative slopes of the curvilinear main effects in our models, therefore, we must plot the interaction terms and also examine the differences in the predicted probabilities of attention allocation associated with different values of the expertise matching variable.

We used the estimates from Model 18 to plot the interaction terms. Since the magnitude, direction, and statistical significance of moderating effects depend on the values of all of the other independent variables in nonlinear models (Hoetker, 2007), and statistical testing of these effects can produce misleading results (Greene, 2010), we follow the suggestion of Greene (2010) and assess the evidence for Hypotheses 4 and 5 by inspecting these plots. To generate the plots, we derived the predicted probabilities of attention allocation at three levels of the moderator variable (expertise matching) over the entire observed range of the moderated variable (e.g., problem length), while holding all other continuous explanatory variables at their means and significant binary variables at 1. We used 1*SD* below and above the mean of the expertise matching variable for the low and high values, respectively, and the mean for the medium value.

The plots for the predicted probabilities for the moderating effects of expertise matching on problem length, breadth, and novelty are presented in Figures 2a, 2b, and 2c. These figures show that a provider was less likely to allocate attention to a problem that was longer, broader, or more novel if the match in expertise between that provider and the problem was low; however, a provider was more likely to allocate attention to such a problem if the level of expertise matching was high (that is, the curves shift upward as expertise matching increases). All three figures also indicate that an increase in expertise matching shifts the maximum of the inverted U-shaped curves toward the right, as illustrated by the vertical dotted lines, suggesting that a closer expertise match increases the point at

[11] Had we found curvilinear-by-linear interactions too, such that $Y = \beta_1 X + \beta_2 X^2 + \beta_3 Z + \beta_4 XZ + \beta_5 X^2 Z$, the derivative, $dY/dX = \beta_1 + 2\beta_2 X + \beta_4 Z + 2\beta_5 XZ$, would have shown that the slope of the curve was a function of X, Z, and $X*Z$. In this case, the slopes of the curve could have changed in additional ways, possibly even to the extent that the inverse U-shaped curve might switch to a U shape for some values of Z. (See, e.g., Van Der Vegt & Bunderson, 2005, in the context of a linear model.)

FIGURE 2
Moderating Effect of Expertise Matching on the Relationship between the Likelihood of Attention Allocation to a Problem and (A) the Length of a Problem (B) the Breadth of a Problem (C) the Novelty of a Problem (D) the number of Competing Problems.



which the costs of allocating attention to a focal problem outweigh the benefits for a knowledge provider.[12]

To further examine whether both the positive slopes and the negative slopes of the curves are significantly affected by expertise matching, we plotted the differences in predicted probabilities associated with a change in expertise matching from low to medium to high values, and then plotted these differences in predicted probabilities against each of the three moderated variables (plots not shown). If expertise matching has the effect of steepening the positive slopes, as well as flattening the negative slopes, of the curves, we would expect to see these differences increase across the entire range of the moderated variables. The plot corresponding to Figure 2a revealed that the differences in predicted probabilities increased to the left of the maximum of the inverted U-shaped curve only for length, indicating that an increase in expertise matching steepened the positive slope of the curve, but did not flatten its negative slope. In contrast, the plots corresponding to Figures 2b and 2c revealed that the differences in predicted probabilities increased across the entire range of the inverted U-shaped curves for both breadth and novelty, indicating that an increase in expertise matching steepened the positive slopes and also flattened the negative slopes of these curves. Thus we find partial support for the moderating effects of expertise matching on problem length predicted in Hypothesis 4a, and full support for the moderating effects of expertise matching on problem breadth and problem novelty predicted in Hypotheses 4b and 4c.

Finally, we plot the predicted probabilities for the moderating effects of expertise matching on the competing problems variable in Figure 2d. This figure shows that providers were more likely to allocate attention to the focal problem when there were more other problems concurrently posted on the forum and that this effect was amplified at higher levels of expertise matching. Deriving and plotting the differences in predicted probabilities for low, medium, and high levels of expertise matching confirmed that the moderating effects of expertise matching were positive across the full range of the competing problems variable.[13] Thus, with the caveat that we did not find a curvilinear main effect for competing problems, Hypothesis 5 is supported, since the effects of competing problems are significantly positively moderated by expertise matching.

Supplementary Analysis

Our hypotheses and empirical analyses focus on a knowledge provider's decision to allocate attention to a particular problem. However, once a knowledge provider has decided to allocate attention to a problem, the amount of time and effort that the provider allocates to that problem may vary—that is, there may be variation in *attention intensity* (Kahneman, 1973; Ocasio, 2011). Our data enable us to examine this in a very preliminary way, by examining the length of the response to a focal problem.

Using the sample of 1,251 problems that received at least one response, we estimated double random effects models that regressed the length of the responses (logged) against the same variables used in our main rare-event logit models.[14] We included the original inverse Mills' ratio derived from the first-stage probit model in which we predicted the likelihood that a provider gave at least one response,

[12] We also derived the confidence intervals for the differences in predicted probabilities using a simulation-based procedure (King, Tomz, & Wittenberg, 2000; Zelner, 2009). Although these confidence intervals need to be interpreted with considerable caution (Greene, 2010), we found that the differences in predicted probabilities associated with a change in expertise matching from low to medium to high values were statistically significant ($p < .05$) for the entire range for each of the three moderated variables (i.e., the confidence intervals around them never contained 0), indicating that the upward shifts of the U-shaped curves were significant. Additionally, we tested whether the rightward shifts of the curves' maxima indicated by the vertical dotted lines were significant. For problem length, we found that the shift in the maximum was significant for the change from low to medium values of expertise matching, but not for the change from medium to high values; for both problem breadth and problem novelty, the shift in the maxima was significant for the changes from low to medium, as well as medium to high, values of expertise matching (i.e., the confidence intervals for these maxima did not overlap).

[13] Again, using a simulation-based procedure to calculate the confidence intervals indicated that this moderating effect of expertise matching was statistically significant.

[14] This specification corrects for the possibility of underestimated standard errors owing to multiple appearances of the same provider and problem in the dataset (see, e.g., Reagans, 2011). Our main rare-event models use clustering instead because an extension of the double random effects approach to such models does not currently exist.

as well as a second inverse Mills' ratio derived from an additional first-stage probit model in which we predicted the likelihood that a problem received at least one response.[15] Thus we controlled for selection bias arising from those individuals who provided at least one response to a problem, as well as from those problems that received at least one response.

Estimates from these models (not shown) indicate that expertise matching had a positive and significant impact on the length of a provider's response to a problem ($\beta = .07$, $p < .01$). The estimates for the linear terms of problem length, breadth, and novelty were all positive and significant (respectively: $\beta = .14$, $p < .01$; $\beta = .10$, $p < .10$; $\beta = .09$, $p < .01$), but the effects for the square terms were negative and significant for problem novelty only ($\beta = -.03$, $p < .10$). The estimates for competing problems were non-significant for either the linear or the squared term. There was evidence of a positive and significant moderating effect of expertise matching for problem length ($\beta = .06$, $p < .05$), but not for any of the other variables. In addition, response order had a positive and significant effect ($\beta = .05$, $p < .01$), indicating that providers gave longer responses to problems that had received more other responses already, and shared projects had a negative and significant effect ($\beta = -.12$, $p < .05$), indicating that providers gave shorter responses to problems posted by seekers with whom they had worked previously. Taken together, these preliminary results suggest that some of the factors that influence the initial decision of *whether* to allocate attention to a problem also influence *how much* attention to allocate subsequently, but the initial decision seems to involve more complex considerations of the benefits and costs of attention allocation; once the commitment is made to allocate some attention to a problem, the costs of allocating more attention seem generally less important.

DISCUSSION

As information demands on managers explode with the growth and spread of social technologies, there is a pressing need for clear explanations of why managers allocate attention to specific problems in digital environments. Our study shifts the scholarly debate from discussions of knowledge provider–seeker relationships (based on relational, social, and reputational rationales) to knowledge provider–problem matches (based on expertise fit, problem characteristics, and problem crowding). Our findings support our central claim that the features of a particular provider–problem match influence attention allocation in an online discussion forum. Below, we address their implications for theories of managerial attention, matching processes, and knowledge sharing in online communities, as well as for our understanding of how social technology platforms are used in organizations.

Attention Allocation as a Matching Process

While prior theories of attention allocation in organizations have offered valuable perspectives on how individuals allocate their attention to problems, they have not focused on how particular individuals allocate attention to particular problems. According to the attention-based view of the firm, for example, the attention of organization members is channeled in some directions and away from others by structural features of organizations such as rules, resources, and relationships (Ocasio, 1997). Relatedly, theories of issue selling emphasize how organization members make deliberate efforts to promote particular problems as worthy of each other's attention (Dutton & Ashford, 1993). Viewing attention allocation as a matching process advances such theories by emphasizing the inherently dyadic nature of this activity and moving beyond a focus on what determines the set of problems that is available for attention allocation to examine how particular individuals allocate attention among the particular problems within that set.

In viewing attention allocation as a matching process between providers and problems, our study is among the first to bring matching theory inside organizations. Originally developed by Becker (1973), matching theory was initially used to explain the formation of marriage partnerships, and subsequently applied to employee–employer matching in labor markets (e.g., Jovanovic, 1979). More recently, it has been extended to an array of matches in interorganizational contexts, including between venture capitalists and startups (Sorensen, 2007), potential alliance partners (Mitsuhashi & Greve, 2009), entrepreneurs and potentially valuable contacts

[15] This second selection model included all of the seeker characteristics in the main outcome models, as well as the variables for problem characteristics and problem crowding. As instrumental variables, we used month and day of the week dummies, based on the assumption that the timing of when a focal problem was posted on the forum would affect the likelihood that it received a response, but not the length of the response that it received.

(Vissa, 2011), and firms and research scientists (Mindruta, 2013). We extend matching theory into the intraorganizational context by examining how matching processes occur within a firm, as part of the daily activities of the organization members. Additionally, while prior research on matching theory has focused on matches between two actors (e.g., employer and employee, or potential alliance partners), we focus on matches between actors and issues—that is, on why individuals allocate attention to particular problems and not others.

One of the core insights of matching theory is that the complementarity between the resources or capabilities of potential partners increases the likelihood of a match (e.g., Mitsuhashi & Greve, 2009; Vissa, 2011). Consistent with this insight, our findings show that greater similarity between the expertise possessed by a provider and the expertise required by the problem increased the likelihood of attention allocation in the online discussion forum that we studied. Furthermore, we found that expertise matching positively moderated the effects of problem length, breadth, and novelty (although it did not increase the likelihood of attention allocation to very long problems). We also found that expertise matching positively moderated the effects of problem crowding, such that an increase in the number of concurrently posted problems was more likely to result in increased attention to the focal problem if the expertise match between the provider and the problem was greater. Thus viewing attention allocation as a matching process leads us toward new ways of understanding why organization members pay attention to some problems and not others.

At What Cost? An Attention Perspective on Knowledge Sharing

Knowledge sharing remains the cornerstone for explanations of how firms leverage the diverse, distributed expertise of their employees to create value and distinguish themselves from competitors (Grant, 1996; Kogut & Zander, 1996). Scholars have made considerable efforts to understand with greater precision how the processes of knowledge sharing unfold within firms (e.g., Argote et al., 2003; Hansen, 1999; Quigley et al., 2007; Reagans & McEvily, 2003; Szulanski, 1996). However, among the broader activities to which organization members can allocate attention, knowledge sharing is often viewed as a peripheral activity (Brown & Duguid, 1991; Lave & Wenger, 1991). This is particularly the case in the context of social technology platforms such as online discussion forums, in which participation is voluntary and often seen as organizational citizenship behavior (Constant et al., 1996; Wasko & Faraj, 2005). In such a context, factors that make it difficult for a knowledge provider to respond to a problem may well crowd out benevolent motivations or the benefits that the provider anticipates from contributing. Our attention perspective on knowledge sharing illuminates such factors by suggesting that knowledge providers take the costs of attention allocation, as well as the benefits, into account in deciding whether or not to respond to particular problems.

In particular, our attention perspective suggests that these costs and benefits will be influenced by the characteristics of a problem itself, as well as by problem crowding. As predicted, our results revealed that problems that were longer, broader, or more novel were more likely to attract attention from a potential knowledge provider—but only up to a point, after which greater length, breadth, or novelty decreased the likelihood of receiving attention. These findings are consistent with our argument that the cognitive load created by a problem that is very long, broad, or novel creates costs for a provider that can outweigh the benefits of these characteristics for attracting attention to the problem.

We expected to find that a higher number of concurrently posted problems would have a similar curvilinear effect on the likelihood of attention allocation to a focal problem, but did not find evidence for this; instead, we found only a positive effect. One possible reason is that the numbers of concurrently posted problems were not high enough in our dataset for a negative effect of competitive crowding to set in. We ran follow-up analyses extending the window for posting other problems from three working days to five, seven, or 10 working days prior to the focal problem, but still found only positive effects. However, when we used the seven- or 10-day windows and also considered only those focal problems with 10 or more competing problems, we found evidence of an inverted U-shaped relationship between the number of competing problems and the likelihood that a provider allocated attention to a focal problem. This suggests that we did not find evidence of such a curvilinear effect in our main models because the maximum value of our concurrent problems variable (12) was below the threshold at which competitive crowding reduces the likelihood that attention is allocated to a focal problem.

One additional provocative finding, although a preliminary one, concerned the effects of provider project load. Contrary to our expectations, the first-stage selection model indicated that individuals who were assigned to more projects in total during the observation period were actually more likely, rather than less likely, to allocate attention to responding to problems on the online discussion forum. This may have been because such individuals were somehow more able or more willing to manage involvement in a wider array of work-related activities. However, consistent with our expectations, our main outcome models showed that a knowledge provider who had a higher project load at the time that a focal problem was posted was less likely to allocate attention to that problem, indicating that the opportunity cost of responding to a problem was higher for such an individual. The implication of these results is that attention allocation is influenced by a provider's attention capacity (see Simon, 1957)—that is, how much attention he or she is able to allocate—in complex ways that are worthy of further exploration in future research.

Taken together, these findings extend theories of knowledge sharing by heeding the call for researchers to pay more "attention to attention" (Ocasio, 2011), and specifically by considering how both the costs and the benefits of allocating attention to a particular problem can influence a potential provider's inclination to share his or her knowledge. While much prior research on knowledge sharing has noted that knowledge seekers face costs, as well as benefits, when trying to secure solutions to their problems through network ties or electronic databases (e.g., Hansen & Haas, 2001; Teece, 1977; Zander & Kogut, 1995), our study breaks new ground by considering the costs, as well as the benefits, that knowledge providers face when allocating their scarce attention to providing such solutions.

Online Knowledge Sharing in Organizations

Our study also aims to contribute to an emerging body of research that specifically focuses on online knowledge sharing in organizations, via social technology platforms such as corporate intranets or databases (e.g., Faraj, Jarvenpaa, & Majchrzak, 2011; Fulk, Heino, Flanagin, Monge, & Bar, 2004; Kankanhalli et al., 2005). In interpersonal contexts, people sometimes choose to withhold their knowledge from others who request it, for practical, strategic, or political reasons (e.g., Connelly, Zweig, Webster, & Trougakos, 2012; Haas & Park, 2010). In online communities, it is even easier to withhold knowledge, since the knowledge seeker does not approach the knowledge provider directly, and thus there is little risk of violating norms or incurring repercussions. For this reason, social technology platforms that are intended to facilitate knowledge sharing are often plagued by collective action problems that deter individuals from contributing their knowledge (e.g., Ba, Stallaert, & Whinston, 2001; Cabrera & Cabrera, 2002; Connolly & Thorn, 1990). Moreover, once they decide to engage in online knowledge sharing, our study shows that knowledge providers make systematic choices about the focus of their contributions that are driven by different considerations from those that drive interpersonal knowledge sharing.

Specifically, much of the increasingly extensive literature on interpersonal knowledge sharing in organizations emphasizes the role of personal connections in facilitating exchanges between individuals, usually through social network ties (e.g., Hansen, 1999; Levin & Cross, 2004; Reagans & McEvily, 2003; Tortoriello & Krackhardt, 2010). However, the control variables in our models indicated that even where actual or potential personal connections between providers and seekers existed, as a result of social similarity, physical proximity, or prior familiarity, these considerations did not increase the likelihood of attention allocation; the only form of connection that mattered in our study was reciprocity. Other studies have found similarly weak evidence for the influence of personal connections in online communities (e.g., Constant et al., 1996). Indeed, the attraction for many organizations of technology platforms such as online discussion forums lies in their ability to facilitate knowledge sharing even in the absence of personal connections between organization members. Yet our understanding of what drives knowledge sharing in such online settings has been limited. In light of this, our study aims to advance research on online knowledge sharing by shifting the focus away from provider–seeker relationships toward provider–problem matching instead, and thus offers insights into how a closer provider–problem expertise match, as well as other characteristics of the problem and problem crowding, influences the likelihood that a provider allocates attention to that problem.

For research on knowledge sharing in organizations, as well as on interpersonal communication and social networks more broadly, there are two notable implications of this shift. First, social network theory has called for more focus on the content of ties, because what is transferred through a tie might influence the choice of partners (Chua, Ingram, & Morris, 2008; Podolny & Baron, 1997). By showing that provider–problem expertise matching influences whether an exchange takes place between a provider and a seeker in an online community, our study heeds this call and highlights the importance of the expertise to be transferred through a tie in determining the activation of that tie. A second implication is that not everything that can be analyzed as a social network necessarily should be analyzed as such. While a provider–problem matrix derived from an online discussion forum can be readily converted into a network of ties between knowledge providers and knowledge seekers, the lack of social context in an online setting limits the fruitfulness of this approach. That said, the more the user interface of an online discussion forum or similar social technology platform is structured in a way that makes social features salient, the more we might expect social network variables to matter for how the platform is used. Thus if an online discussion forum were to be designed in a way that makes the characteristics of its knowledge seekers highly salient to its knowledge providers, for example by requiring seekers to post their photos or locations with their questions, factors such as social similarity or physical proximity might drive knowledge sharing more than we observed in a setting in which these characteristics were not highly salient.

Future Directions

Our study of knowledge sharing in an online discussion forum illuminates how knowledge providers decide whether or not to allocate their attention to particular problems. The study has its limitations, however, which suggest some potential avenues for future research. The first relates to the generalizability of our results, given that we focus on a single professional services organization. Although we have a large sample of individuals, and a considerable amount of information about them, we must look to future research to establish the extent to which our findings reflect the particular features of the organization, or alternatively reveal more general patterns. For example, the firm studied has many employees dispersed around the world and a large number of online discussion forums. Thus it could be that this is an organization in which employees are more selective about which problems they choose to address than might be the case, for example, in a smaller organization in which there is more pressure to participate, in which an online discussion forum is a relatively novel and exciting technology, or in which contributing knowledge by responding to problems is viewed as a way in which to signal status.

Second, in focusing on the matching process between knowledge providers and problems, we limited our scope to studying whether a provider posted a response to a problem on the online discussion forum. In our supplementary analysis, we also examined how much attention they allocated, as measured by the length of their response. However, we recognize that this supplementary analysis is more suggestive than definitive, since longer responses may or may not actually take more time and effort to formulate than shorter answers. Using additional measures and exploring the distinctive drivers of attention intensity more fully would therefore be a valuable direction for future research. Moreover, our data did not allow us to evaluate the quality of the responses provided to a problem. Further research could usefully examine the impact of provider–problem expertise matching, problem characteristics, and problem crowding on the quality of online knowledge sharing, perhaps in a research setting in which knowledge providers are rated by their colleagues on the helpfulness of their online contributions. Finally, in focusing on why providers allocate attention to particular problems in an online discussion forum, we have not addressed the question of why knowledge seekers post problems to the forum in the first place, whether certain types of individual are more likely to post problems than others, or whether certain types of problem are more likely to be posted. These would also be useful directions for future research.

In conclusion, this study offers fresh insights into online knowledge sharing in organizations by examining why individuals choose to allocate attention to specific problems. From a knowledge management perspective, social technology platforms such as online discussion forums are valuable tools for facilitating knowledge sharing among globally dispersed employees. However, the ability of organizations to realize the full potential of these tools is limited by the attention that their members choose to devote to providing solutions to each other's problems. By viewing attention allocation in an online discussion forum as a matching process between providers and problems, this study adds to current debates on how knowledge is shared within organizations, especially in the increasingly important online context.

REFERENCES

Ahuja, M. K., & Carley, K. M. 1999. Network structure in virtual organizations. ***Organization Science***, 10: 741–757.

Ahuja, M. K., Galletta, D. F., & Carley, K. M. 2003. Individual centrality and performance in virtual R&D groups: An empirical study. ***Management Science***, 49: 21–38.

Aiken, L. S., & West, S. G. 1991. ***Multiple regression: Testing and interpreting interactions***. London: Sage.

Allen, T. 1977. ***Managing the flow of technology***. Cambridge, MA: MIT Press.

Aral, S., & Van Alstyne, M. 2011. The diversity–bandwidth trade-off. ***American Journal of Sociology***, 117: 90–171.

Argote, L., McEvily, B., & Reagans, R. 2003. Managing knowledge in organizations: An integrative framework and review of emerging themes. ***Management Science***, 49: 571–582.

Ba, S., Stallaert, J., & Whinston, A. B. 2001. Optimal investment in knowledge within a firm using a market mechanism. ***Management Science***, 47: 1203–1219.

Becker, G. S. 1973. A theory of marriage, part I. ***Journal of Political Economy***, 81: 813–846.

Bock, G. W., Zmud, R. W., Kim, Y. G., & Lee, J. N. 2005. Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. ***Management Information Systems Quarterly***, 29: 87–111.

Brown, J. S., & Duguid, P. 1991. Organizational learning and communities-of-practice: Towards a unified view of working, learning and innovation. ***Organization Science***, 2: 40–57.

Brown, J. S., & Duguid, P. 2002. ***The social life of information***, Cambridge, MA: Harvard Business School Press.

Burgelman, R. A. 1991. Intraorganizational ecology of strategy-making and organization adaptation: Theory and field research. ***Organization Science***, 2: 239–262.

Butler, B. S. 2001. Membership size, communication activity, and sustainability: A resource-based model of online social structures. ***Information Systems Research***, 12: 346–362.

Cabrera, A., & Cabrera, E. F. 2002. Knowledge-sharing dilemmas. ***Organization Studies***, 23: 687–710.

Chandler, P., & Sweller, J. 1991. Cognitive load theory and the format of instruction. ***Cognition and Instruction***, 8: 293–332.

Chiu, C. M., Hsu, M. H., & Wang, E. T. G. 2006. Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. ***Decision Support Systems***, 42: 1872–1888.

Chua, R., Ingram, P., & Morris, M. 2008. From the head and the heart: Locating cognition- and affect-based trust in managers' professional networks. ***Academy of Management Journal***, 51: 436–452.

Cohen, M. D., March, J. G., & Olsen, J. P. 1972. A garbage can model of organizational choice. ***Administrative Science Quarterly***, 17: 1–25.

Cohen, W. M., & Levinthal, D. A. 1990. Absorptive capacity: A new perspective of learning and innovation. ***Administrative Science Quarterly***, 35: 128–152.

Connelly, C. E., Zweig, D., Webster, J., & Trougakos, J. P. 2012. Knowledge hiding in organizations. ***Journal of Organizational Behavior***, 33: 64–88.

Connolly, T., & Thorn, B. K. 1990. Discretionary databases: Theory, data, and implications. In J. Fulk & C. W. Steinfield (Eds.), ***Organizations and communication technology***: 219–233. Newbury Park, CA: Sage.

Constant, D., Sproull, L., & Kiesler, S. 1996. The kindness of strangers: The usefulness of electronic weak ties for technical advice. ***Organization Science***, 7: 119–135.

Corrocher, N., Malerba, F., & Montobbio, F. 2007. Schumpeterian patterns of innovative activity in the ICT field. ***Research Policy***, 36: 418–432.

Cramton, C. D. 2001. The mutual knowledge problem and its consequences for dispersed collaboration. ***Organization Science***, 12: 346–371.

Criscuolo, P., Salter, A., & Sheehan, T. 2007. Making knowledge visible: Using expert yellow pages to map capabilities in professional services firms. ***Research Policy***, 36: 1603–1619.

Cummings, J. N. 2004. Work groups, structural diversity, and knowledge sharing in a global organization. ***Management Science***, 50: 352–364.

Cyert, R. M., & March, J. G. 1963. ***A behavioural theory of the firm***. Englewood Cliffs, NJ: Prentice-Hall.

Dane, E. 2010. Reconsidering the trade-off between expertise and flexibility: A cognitive entrenchment perspective. ***Academy of Management Review***, 35: 579–603.

Davenport, T. H., & Beck, J. C. 2001. ***The attention economy: Understanding the new currency of business***. Cambridge, MA: Harvard Business School.

Davenport, T. H., & Prusak, L. 2000. ***Working knowledge: How organizations manage what they know***. Cambridge, MA: Harvard Business School.

Dean, D., & Webb, C. 2011. Recovering from information overload. ***The McKinsey Quarterly***, 1: 80–88.

Duda, R. O., & Hart, P. E. 1973. ***Pattern recognition and scene analysis***. New York: Wiley.

Dudley, N. M., & Cortina, J. M. 2008. Knowledge and skills that facilitate the personal support dimension of citizenship. ***The Journal of Applied Psychology***, 93: 1244–1270.

Dutton, J. E., & Ashford, S. J. 1993. Selling issues to top management. ***Academy of Management Review***, 18: 397–428.

Dutton, J. E., Ashford, S. J., O'Neill, R. M., & Lawrence, K. A. 2001. Moves that matter: Issue selling and organizational change. ***Academy of Management Journal***, 44: 716–736.

Eggers, J. P., & Kaplan, S. 2009. Cognition and renewal: Comparing CEO and organizational effects on incumbent adaptation to technical change. ***Organization Science***, 20: 461–477.

Espinosa, J. A., Slaughter, S. A., Kraut, R. E., & Herbsleb, J. D. 2007. Familiarity, complexity, and team performance in geographically distributed software development. ***Organization Science***, 18: 613–630.

Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. 2011. Knowledge collaboration in online communities. ***Organization Science***, 22: 1224–1239.

Fulk, J., Heino, R., Flanagin, A. J., Monge, P. R., & Bar, F. 2004. A test of the individual action model for organizational information commons. ***Organization Science***, 15: 569–585.

Gardner, H. 1985. ***Frames of mind: The theory of multiple intelligences***. New York: Basic Books.

George, G., Kotha, R., & Zheng, Y. 2008. The puzzle of insular domains: A longitudinal study of knowledge structuration and innovation in biotechnology firms. ***Journal of Management Studies***, 45: 1448–1474.

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

Greene, W. 2010. Testing hypotheses about interaction terms in nonlinear models. ***Economics Letters***, 107: 291–296.

Gulati, R. 1995. Social structure and alliance formation patterns: A longitudinal analysis. ***Administrative Science Quarterly***, 40: 619–652.

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

Hallen, B. L. 2008. The causes and consequences of the initial network positions of new organizations: From whom do entrepreneurs receive investments? ***Administrative Science Quarterly***, 53: 685–718.

Hansen, M. T. 1999. The search–transfer problem: The role of weak ties in sharing knowledge across organizational subunits. ***Administrative Science Quarterly***, 44: 82–111.

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

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

Heckman, J. J. 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. ***Annals of Economic and Social Measurement***, 5: 475–492.

Higgins, E. T. 1996. Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), ***Social psychology: Handbook of basic principles***: 133–168. New York: Guilford Press.

Hoetker, G. 2007. The use of logit and probit models in strategic management research: Critical issues. ***Strategic Management Journal***, 28: 331–343.

Hsu, M. H., Ju, T. L., Yen, C. H., & Chang, C. M. 2007. Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations. ***International Journal of Human–Computer Studies***, 65: 153–169.

James, W. 1890. ***The principles of psychology***. New York: Dover Publications.

Jensen, M. 2003. The role of network resources in market entry: Commercial banks' entry into investment banking, 1991–1997. ***Administrative Science Quarterly***, 48: 466–497.

Jeppesen, L. B., & Frederiksen, L. 2006. Why do users contribute to firm-hosted user communities? The case of computer-controlled music instruments. ***Organization Science***, 17: 45–63.

Jeppesen, L. B., & Lakhani, K. R. 2010. Marginality and problem-solving effectiveness in broadcast search. ***Organization Science***, 21: 1016–1033.

Jones, Q., David, G., & Rafaeli, S. 2004. Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. ***Information Systems Research***, 15: 194–210.

Joseph, J., & Ocasio, W. 2012. Architecture, attention, and adaptation in the multibusiness firm: General electric from 1951 to 2001. ***Strategic Management Journal***, 33: 633–660.

Jovanovic, B. 1979. Job matching and the theory of turnover. ***Journal of Political Economy***, 87: 972–990.

Kahneman, D. 1973. ***Attention and effort***. Englewood Cliffs, NJ: Prentice-Hall.

Kane, G. C., & Alavi, M. 2007. Information technology and organizational learning: An investigation of exploration and exploitation processes. ***Organization Science***, 18: 796–812.

Kankanhalli, A., Tan, B. C. Y., & Wei, K. K. 2005. Contributing knowledge to electronic knowledge repositories: An empirical investigation. ***Management Information Systems Quarterly***, 29: 113–143.

King, G., & Zeng, L. 2001. Logistic regression in rare events data. ***Political Analysis***, 9: 137–163.

King, G., Tomz, M., & Wittenberg, J. 2000. Making the most of statistical analyses: Improving interpretation and presentation. ***American Journal of Political Science***, 44: 347–355.

Kogut, B., & Zander, U. 1996. What firms do? Coordination, identity, and learning. ***Organization Science***, 7: 502–518.

Kotha, R., George, G., & Srikanth, K. 2013. Bridging the mutual knowledge gap: Coordination costs and the commercialization of science. ***Academy of Management Journal***, 56: 498–524.

Krackhardt, D. 1988. Predicting with networks: Nonparametric multiple-regression analysis of dyadic data. ***Social Networks***, 10: 359–381.

Lakhani, K. R., & Wolf, R. 2005. Why hackers do what they do: Understanding motivation and effort in free/open source software projects. In J. Feller, B. Fitzgerald, S. Hissam & K. R. Lakhani (Eds.), ***Perspectives on free and open-source software***. Cambridge, MA: MIT Press.

Lave, J., & Wenger, E. 1991. ***Situated learning: Legitimate peripheral participation***. Cambridge: Cambridge University Press.

Levin, D. Z., & Cross, R. 2004. The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. ***Management Science***, 50: 1477–1490.

Levinthal, D., & Rerup, C. 2006. Crossing an apparent chasm: Bridging mindful and less-mindful perspectives on organizational learning. ***Organization Science***, 17: 502–513.

Li, Q., Maggitti, P., Smith, K., Tesluk, P., & Katila, R. 2013. Top management attention to innovation: The role of search selection and intensity in new product introductions. ***Academy of Management Journal***, 56: 893–916.

Lin, M. J., Hung, J., & Chen, C. 2009. Fostering the determinants of knowledge sharing in professional virtual communities. ***Computers in Human Behavior***, 25: 929–939.

Ma, M., & Agarwal, R. 2007. Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. ***Information Systems Research***, 18: 42–67.

March, J. G. 1994. ***Primer on decision making: How decisions happen***. New York: The Free Press.

March, J. G. 2002. Understanding how decisions happen in organizations. In Z. Shapira (Ed.), ***Organizational decision making***: 9–34. Cambridge: Cambridge University Press.

March, J. G., & Olsen, J. P. 1976. Organizational learning under ambiguity. ***European Journal of Political Research***, 3: 147–171.

March, J. G., & Simon, H. A. 1958. ***Organizations***. New York: John Wiley & Sons.

Markus, M. L. 1987. Toward a "critical mass" theory of interactive media universal access, interdependence and diffusion. ***Communication Research***, 14: 491–511.

McArthur, L. Z. 1981. What grabs you? The role of attention in impression formation and causal attribution. In E. T., C. P. Herman & M. P. Zanna (Eds.), ***Social cognition: The Ontario Symposium, vol. 1***: 201–241. Hillsdale, NJ: Erlbaum.

McKinsey Global Institute. 2012. ***The social economy: Unlocking value and productivity through social technologies***. Washington, DC: McKinsey & Co.

McPherson, M., Smith-Lovin, L., & Cook, J. M. 2001. Birds of a feather flock together: Homophily in social networks. ***Annual Review of Sociology***, 27: 415–444.

Mindruta, D. 2013. Value creation in university–firm research collaborations: A matching approach. ***Strategic Management Journal***, 34: 644–665.

Mitsuhashi, H., & Greve, H. R. 2009. A matching theory of alliance formation and organizational success: Complementarity and compatibility. ***Academy of Management Journal***, 52: 975–995.

Nahapiet, J., & Ghoshal, S. 1998. Social capital, intellectual capital and the organisational advantage. ***Academy of Management Review***, 23: 242–266.

Nelson, D. L. 1979. Remembering pictures and words: Appearance, significance, and name. In L. S. Cermak & E. I. M. Craik (Eds.), ***Levels of processing in human memory***: 45–76. Hillsdale, NJ: Erlbaum.

Neter, J., Wasserman, W., & Kutner, M. H. 1990. ***Applied linear statistical models***. Homewood, IL: Irwin.

Ocasio, W. 1997. Towards an attention-based view of the firm. ***Strategic Management Journal***, 18: 187–206.

Ocasio, W. 2011. Attention to attention. ***Organization Science***, 22: 1286–1296.

Paas, F., van Gog, T., & Sweller, J. 2010. Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. ***Educational Psychology Review***, 22: 115–121.

Parkhurst, D., Law, K., & Niebur, E. 2002. Modeling the role of salience in the allocation of overt visual attention. ***Vision Research***, 42: 107–123.

Piezunka, H., & Dahlander, L. 2014. Distant search, narrow attention: How crowding alters organizations' filtering of suggestions in crowdsourcing. ***Academy of Management Journal***, Advance online publication [DOI: 10.5465/amj.2012.0458].

Podolny, J. M., & Baron, J. N. 1997. Resources and relationships: Social networks and mobility in the workplace. ***American Sociological Review***, 62: 673–693.

Quigley, N. R., Tesluk, P. E., Locke, E. A., & Bartol, K. M. 2007. A multilevel investigation of the motivational mechanisms underlying knowledge sharing and performance. ***Organization Science***, 18: 71–88.

Radicati. 2013. ***Email statistics report 2013–2017***. Palo Alto, CA: Radicati Group.

Reagans, R. 2005. Preferences, identity, and competition: Predicting tie strength from demographic data. ***Management Science***, 51: 1374–1383.

Reagans, R. 2011. Close encounters: Analyzing how social similarity and propinquity contribute to strong network connections. ***Organization Science***, 22: 835–849.

Reagans, R., & McEvily, B. 2003. Network structure and knowledge transfer: The effects of cohesion and range. ***Administrative Science Quarterly***, 48: 240–267.

Ross, L., & Nisbett, R. E. 1991. ***The person and the situation: Perspectives of social psychology***. New York: McGraw-Hill.

Sambamurthy, V., & Subramani, M. 2005. Special issue on information technologies and knowledge management. ***Management Information Systems Quarterly***, 29: 193–195.

Shapiro, C., & Varian, H. R. 1999. ***Information rules: A strategic guide to the network economy***. Boston, MA: Harvard Business School Press.

Simon, H. A. 1947. ***Administrative behavior: A study of decision-making process in administrative organization***. Chicago, IL: Macmillan.

Simon, H. A. 1957. ***Models of man: Social and rational***. New York: John Wiley & Sons.

Sole, D., & Edmondson, A. 2002. Situated knowledge and learning in dispersed teams. ***British Journal of Management***, 13: S17–S34.

Sorensen, M. 2007. How smart is smart money? A two-sided matching model of venture capital. ***The Journal of Finance***, 62: 2725–2762.

Sullivan, B. N. 2010. Competition and beyond: Problems and attention allocation in the organizational rule-making process. ***Organization Science***, 21: 432–450.

Sweller, J. 1988. Cognitive load during problem solving: Effects on learning. ***Cognitive Science***, 12: 257–285.

Szulanski, G. 1996. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. ***Strategic Management Journal***, 17: 27–43.

Taylor, S. E., & Fiske, S. T. 1978. Advances in experimental social psychology. In L. Berkowitz (Ed.), ***Salience, attention, and attribution: Top of the head phenomena, vol. 11***: 249–288. New York: Academic Press.

Teece, D. 1977. Technology transfer by multinational firms: The resource cost of transferring technological know-how. ***The Economic Journal***, 87: 242–261.

Thomas, J. B., Sussman, S. W., & Henderson, J. C. 2001. Understanding "strategic learning": Linking organizational learning, knowledge management, and sense-making. ***Organization Science***, 12: 331–345.

Tomz, M. 2003. Relogit: Rare events logistic regression. ***Journal of Statistical Software***, 8: 137–163.

Tortoriello, M., & Krackhardt, D. 2010. Activating cross-boundary knowledge: The role of simmelian ties in the generation of innovations. ***Academy of Management Journal***, 53: 167–181.

Tsai, W. 2001. Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. ***Academy of Management Journal***, 44: 996–1004.

Van Der Vegt, G. S., & Bunderson, J. S. 2005. Learning and performance in multidisciplinary teams: The importance of collective team identification. ***Academy of Management Journal***, 48: 532–547.

Vissa, B. 2011. A matching theory of entrepreneurs' tie formation intentions and initiation of economic exchange. ***Academy of Management Journal***, 54: 137–158.

Walther, J. B., Anderson, J. F., & Park, D. W. 1994. Interpersonal effects in computer-mediated interaction a meta-analysis of social and antisocial communication. ***Communication Research***, 21: 460–487.

Wasko, M., & Faraj, S. 2000. It is what ones does: Why people participate and help others in electronic communities of practice. ***The Journal of Strategic Information Systems***, 9: 155–173.

Wasko, M., & Faraj, S. 2005. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. ***Management Information Systems Quarterly***, 29: 35–57.

Weick, K. E. 1979. ***The social psychology of organizing***. Reading, MA: Addison-Wesley.

Weick, K. E., & Sutcliffe, K. M. 2006. Mindfulness and the quality of organizational attention. ***Organization Science***, 17: 514–524.

Wolfe, J. M., & Horowitz, T. S. 2004. What attributes guide the deployment of visual attention and how do they do it? ***Nature Reviews: Neuroscience***, 5: 495–501.

Zahra, S. A., & George, G. 2002. Absorptive capacity: A review, reconceptualisation, and extension. ***Academy of Management Review***, 27: 185–203.

Zander, U., & Kogut, B. 1995. Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test. ***Organization Science***, 6: 76–92.

Zelner, B. A. 2009. Using simulation to interpret results from logit, probit, and other nonlinear models. ***Strategic Management Journal***, 30: 1335–1348.

Paola Criscuolo (p.criscuolo@imperial.ac.uk) is an associate professor of innovation management at Imperial College Business School. She received her PhD in the economics of innovation and technological change from Maastricht University. Her research interests include knowledge transfer and innovation at the firm and individual levels.

Gerry George (ggeorge@smu.edu.sg) is dean and professor of innovation and entrepreneurship at Lee Kong Chian School of Business at Singapore Management University. He received his PhD from Virginia Commonwealth University. His recent work examines strategic and organizational factors, such as business models and organizational design, and their impact on innovation and performance.

Martine R. Haas (mrhaas@wharton.upenn.edu) is an associate professor of management at the Wharton School, University of Pennsylvania. She received her PhD in organizational behavior from Harvard University. Her research focuses on collaboration in global organizations, with an emphasis on knowledge sharing and team effectiveness.