Crowd Experience and Performance:
An Empirical Analysis of Crowdsourced New Product Development

Completed Research Paper

Zhiyi Wang
National University of Singapore
13 Computing Drive
Singapore 117417
wangzhiyi@nus.edu.sg

Jungpil Hahn
National University of Singapore
13 Computing Drive
Singapore 117417
jungpil@nus.edu.sg

Abstract

Crowdsourcing has widely been used as a strategy for sourcing ideas and efforts to facilitate innovation and new product development. However, research into the value creation of crowdsourcing and the efficacy of the crowd is still limited. In this study, we investigate a crowdsourced new product development context and apply the framework of collective intelligence to understand the crowd co-creation process. Based on the theory of diversity about generalists vs. specialists, we examine the role of prior experience in affecting crowd performance. Our empirical analysis shows that participants with both diverse and specialized experience are helpful in enhancing crowd performance in terms of efficient product development. We also find participants with T-shaped experience in non-focal tasks are beneficial. Contrary to other group contexts, generalists do not seem to be a valuable type in our study context. The findings provide insights for understanding collaboration and value co-creation in crowdsourcing communities.

Keywords: Crowdsourcing, Diversity, Generalists, Specialists, Collective Intelligence, New Product Development
Introduction

Creating innovations is no longer the sole purview of domain experts but has recently become accessible to ordinary people. The “wisdom of crowds”, or collective intelligence, has been appropriated by multiple stakeholders including firms, governments, scientists, technical experts as well as researchers (Howe 2006; Howe 2008; Malone et al. 2010; Savage 2012). Numerous crowd-based platforms have been established to facilitate innovations and create new forms of economic value (Avital et al. 2014). Crowdsourcing, a term coined by Jeff Howe in 2006 (Howe 2006), has been widely used for seeking various information and knowledge in a variety of domains. Platforms such as IdeaStorm, TopCoder and IndieGoGo have attracted numerous users who are not formal domain experts but nonetheless contribute important assets (e.g., ideas, knowledge or money) to problem solving and innovation. Leading companies including Dell, Starbucks, SAP, GE and Apple have also built crowdsourcing platforms and communities to attract value creation from the crowd (Krcmar et al. 2009; Ramaswamy and Gouillart 2010).

Along with the popularity of crowdsourcing and crowd-based platforms, a variety of business models have been developed by companies and studied by scholars. Traditional crowdsourcing typically takes the form of a one-time collection of ideas or “contests”, while Web 2.0 and social media introduced social platforms and communities for crowdsourcing. Many crowdsourcing platforms adopt a competition model and participants compete against one another to win the contest (Terwiesch and Xu 2008). Another form of crowdsourcing is firm-oriented, a model whereby a company hosts a platform and community members contribute new ideas to company’s services and R&D. IdeaStorm, the well-known platform hosted by Dell, is a representative example of this model (Bayus 2013).

Although competition and idea generation have been regarded as dominating forms of crowdsourcing, novel elements have been introduced. One of the new emergent elements in the crowdsourcing business model, which is the focus of this study, is collaboration. In contrast to the competition model where participants do not interact with one another, this business model emphasizes the notion of “crowd co-creation” and introduces interdependencies in crowdsourcing projects (Avital et al. 2014; Malone et al. 2010). A good example of such a crowd co-creation process is the LEGO Ideas Community, which leads user collaboration for value co-creation (Antorini et al. 2012). The penetration of social media, online communities and digital communication enables such collaboration-based crowdsourcing to be practically feasible. Compared to traditional crowdsourcing of work using contests and ideation forums, the concept of value co-creation is more salient in the collaboration process, which could lead to more predictable value propositions from the crowd and a more sustainable community (Nguyen et al. 2013; Nickerson et al. 2014). Understanding the crowd-based co-creation process would help to generate new insights into the new emerging concepts such as the “crowd-based economy” (Avital et al. 2014) and “Innovation 2.0” (Marjanovic et al. 2012). Innovations from outside have evolved into new forms (Song et al. 2009) and have become more democratized by relying on the “Inventor” or “Maker” groups that are dedicated to make the ideas come true (Anderson 2014). Such new emerged concepts and phenomena in crowdsourcing and innovation have altered how crowds work and how value is created in these communities. However, despite the growing popularity of collaboration elements and co-creation processes in crowdsourcing models, research has yet to thoroughly investigate them. The existing literature has primarily focused on competitive contest-based and idea generation processes in value creation in crowdsourcing; only limited attention has been paid to new forms of innovation in crowdsourcing communities (Dissanayake et al. 2014), especially collaborative intelligence and crowd co-creation process in crowdsourcing. Given the growing importance of the collaboration economy and value co-creation, it is important to investigate these novel phenomena in crowdsourcing to better understand the value creation process and add new insights for the crowdsourcing literature.

In this study, we investigate crowdsourced new product development (which adopts both collection and collaboration for collective intelligence) using the framework of collective intelligence (Malone et al. 2010) in virtual communities. To understand how the crowd works and creates value, we adopt the theoretical lens from the diversity literature to explain different outcomes derived by the crowd in new product development. More specifically, we consider the role of prior experience of crowd participants in affecting crowdsourcing performance. We examine the following research question: how do prior experiences of crowd members impact crowd performance in crowdsourced new product development? Three reasons motivate us to investigate this research question. First, although the role of experience has been examined...
in the crowdsourcing literature (Archak and Ghose 2010; Huang et al. 2012), competing results exist and the characteristics (e.g., breadth and depth) of prior experience are still not well-understood. Participants with different characteristics of experience will have different learning outcomes and contribute heterogeneously to the crowdsourced work. We try to fill this gap by examining the types of participants in the crowd based on characteristics of participants’ prior experiences. In addition, as the performance of crowdsourcing has traditionally been measured by prior research through peer ratings and extent of participation (Hwang et al. 2014; Yang et al. 2009), our knowledge of how crowdfunders perceive the outcome of crowdsourcing is still limited. The existing literature mainly focused on participation and ideation processes rather than on the value co-created. Also, the literature neglects how collectively the crowd works and performs (Pedersen et al. 2013). Finally, the theory about diversity and composition in creative teams or groups has traditionally focused on small groups (Taylor and Greve 2006). Whether theoretical insights generalize to large and disperse online groups (i.e., the crowd) (Majchrzak and Malhotra 2013) should be ascertained. The new business model of crowdsourcing and value co-creation provides a context for us to investigate related theories in the large and disperse groups. Therefore, answering our research question will help to better understand the new crowd-based economy and value co-creation processes in crowdsourcing communities.

To examine the various types of the crowd members based on their experience, we develop a typology based on the framework of generalist vs. specialist and hypotheses from the diversity literature about teams and groups. Data from Quirky.com on new product development allows us to empirically identify six types of crowd participants. Furthermore, we find that a crowd with a greater proportion of members who possess experiences that are both high in diversity and in specialization is associated with better performance in terms of product development efficiency. In addition, a greater proportion of participants with a T-shaped experience distribution with specialization in non-focal tasks in the crowd was also found to positively affect performance. Finally, we find that generalists are not an ideal type in our context. By capturing the dynamics of crowd construction in the crowdsourcing context and examining the characteristics of participant experiences, our study contributes to the literatures on crowdsourcing, open innovation, virtual communities, as well as to the theory of diversity and group/team composition.

**Literature Review and Theoretical Background**

**Crowdsourcing**

Crowdsourcing has been studied by many scholars in recent years. Jeff Howe coined the term to denote the outsourcing of internal tasks to outside individuals through an open call (Howe 2006; Howe 2008). Crowdsourcing typically calls for the collective intelligence to facilitate the process of new product development, business analytics and problem solving. However, different types of crowdsourcing exist and have been investigated by different research streams (Huang et al. 2012; Moqri et al. 2014). Here we review the key elements that characterize major types of crowdsourcing (Geiger et al. 2011; Pedersen et al. 2013): competition (e.g., TopCoder), collaboration (e.g., Lego) and idea generation (e.g., IdeaStorm and Giffgaff). These elements are consistent with the recent research agenda for crowdsourcing (Estelles-Arolas and Gonzalez-Ladron-de-Guevara 2012; Majchrzak and Malhotra 2013).

In the first element (i.e., competition), crowdsourcing participants provide their solutions to a specific task such as a code segment or logo design, and only those who submit the best works can be selected as winners and earn a reward. Competition-based crowdsourcing is typically regarded as auction or contest and economic models were used to investigate how participants behave in the contest (Archak and Ghose 2010; Archak and Sundararajan 2009; Huang et al. 2012; Koh 2014; Yang et al. 2009). Factors affecting crowdsourcing quality (Krmar et al. 2009; Shao et al. 2012; Yang et al. 2009) and motivation (Hou et al. 2011; Jiang and Wagner 2014) were also examined. In these studies, how to attract more problem solvers (project level) and how to win the contest (individual level) are the focal questions examined.

For the second element (i.e., collaboration), participants in the crowd collaboratively work on a collective outcome. Participants may collaborate as in a virtual team (e.g., in open source software development). They may also work like a larger and sparse group in both collective and collaborative ways. Conceptual models including the peer-production model (Haythornthwaite 2009; Nguyen et al. 2013) and patterns of collaboration (Nguyen et al. 2013; Vreede et al. 2009) have been used to provide insights into the collaboration process. Paulini et al. (2013) investigated the collaborative design process in innovation
communities through qualitative analysis of the communication in forums. However, discussion about this element in crowdsourcing is still limited in the literature. More empirical examinations are required to better understand the collaborative component of crowdsourcing. Our study focuses on a model with this element empirically and tries to fill this gap.

The third element (i.e., idea generation) is typically adopted by firms to crowdsource new product ideas (Bayus 2013; Ramaswamy and Goullart 2010) and typically does not require a specific task for the crowd. Instead, through an open call, firms benefit from new innovative ideas (Krcmar et al. 2009) and consumers also take the opportunity to make their ideas come to life. Studies on IdeaStorm investigated the negative effects of ideators’ past success on their performance (Bayus 2013) and consumers’ self-updating about their own ability and firm’s implementation cost (Huang et al. 2014). Hwang et al. (2014) found that an individual’s performance in idea generation is influenced by her knowledge. The adoption and quality of ideas are the focal interest for this element in literature.

This study examines a crowdsourcing model with the collaboration element and can be conceptualized as a group of individuals (i.e., the crowd) with dynamic participation. We adopt the framework of “collective intelligence” (Malone et al. 2010) to provide foundations for our conceptualization. On the crowd level, we investigate how the crowd members’ prior experiences play a role in the performance of the dynamically formed crowd (Ebel et al. 2014). Next, we review related literature about experience in crowdsourcing.

The Role of Experience in Crowdsourcing

A series of studies have shown the effects of experience in crowdsourcing. Most of them focused on the effects of learning from prior experience on individual’s behavior in crowdsourcing contests. Archak and Ghose (2010) examined the learning-by-doing effects at TopCoder and they found that coders would both myopically learn from their experience through participating in the same programming language projects and try other new language projects in a forward-looking manner. Huang et al. (2012) also examined the experience-based learning at Threadless, showing that experience reduces the effort in submitting solutions. However, past experience does not always improve individual’s performance and efforts. Successful experience in idea generation was shown to cause cognitive fixation on ideators and negatively affect their subsequent performance (Bayus 2013). In addition, other outcomes such as subsequent participation and strategic behaviors were also investigated (Huang et al. 2014; Yang et al. 2008).

The existing literature provides evidence that prior experience does matter in crowdsourcing, but competing effects are discussed and proposed. Based on this, our study investigates the characteristics of experiences of crowd members through the framework of generalists vs. specialists, based on their experience in past crowdsourcing tasks. Next, we review the literature comparing generalist and specialist.

Generalist vs. Specialist

Generalist and specialist are generally defined by the distribution of their knowledge (Rulke and Galaskiewicz 2000). Based on the portfolio of knowledge and experience, generalist and specialist can be determined by how diverse/specialized their experience and knowledge are (Boh et al. 2007; Kang et al. 2012; Narayanan et al. 2009). According to the diversity literature, generalists usually possess higher knowledge diversity (breath), while specialists possess deep knowledge with restricted domains. Thus, a generalist can be defined as someone who has experience/knowledge in a broad or diverse range of domains but does not have deep experience/knowledge in any particular area. On the other hand, a specialist can be defined as someone who has highly specialized experience in some focal areas but does not have enough experience in a broad range of knowledge domains.

It has been widely shown that generalists (with greater knowledge diversity) could enhance performance and productivity. Greater knowledge diversity is regarded as a source of creativity (Taylor and Greve 2006), and generalists have the ability to learn new knowledge more efficiently. They know how to learn new knowledge and skills with reduced errors by better relating to their existing stock of knowledge based on their highly diverse experiences (Narayanan et al. 2009). They are also able to easily search from existing solutions to address new tasks (Narayanan et al. 2009). In addition, Hwang et al. (2014) proposed that broad knowledge could induce novelty (i.e., think outside the box). Bayus (2013) found that the diverse experience in commenting activities could mitigate the cognitive fixation problem in idea generation. Also, from a group perspective, Rulke and Galaskiewicz (2000) found that having more
generalists in a group could enhance group performance by facilitating knowledge exchange and sharing. These findings imply that diverse experience (i.e., generalists) should help to enhance performance.

Conversely, some studies caution that the diversity of experience may hinder performance. The variety of experience has an inverted U-shaped effect on software maintenance productivity, which means that too much diversity harms productivity (Narayanan et al. 2009). Also, the effect of experience variety was shown to be moderated by tasks relatedness (i.e., conflict between the new knowledge and past experiences) (Armstrong and Hardgrave 2007; Boh et al. 2007). Thus, it is not always the case that diversity of experience improves performance.

On the other hand, there is also evidence suggesting that specialists (with deep knowledge in limited domains instead of a bit of knowledge in diverse domains) can enhance performance and productivity. The literature in software development has shown that specialized experience enhances team learning and productivity (Boh et al. 2007; Huckman et al. 2009; Kang et al. 2012; Narayanan et al. 2009). Hwang et al. (2014) also showed that in the innovation context, generalists with deep knowledge in at least one domain could generate better ideas than those with shallow knowledge, emphasizing the importance of knowledge depth in innovation. However, other studies also showed that specialized experience is not always beneficial for performance and productivity. Specialization in some knowledge types may cause learning myopia and undermine the ability to learn new knowledge (Archak and Ghose 2010). In addition, the effect of past experience is marginally decreasing and becomes less significant after a certain point (Argote 2012). Cognitive fixation may also occur as specialized experiences increase (Bayus 2013). Thus, there are still conflicting results regarding the role of specialists and specialized experiences.

Given the conflicting results in the literature, we develop a typology of different types of crowd members based on generalist (diverse experience) and specialist (specialized experience) and use this to develop hypotheses using the mechanisms in the diversity literature. As most of the literature examines team diversity and composition (i.e., generalist and specialist) in the context of small groups, our study also tries to investigate the phenomenon in the context of large and disperse groups.

**Collective Intelligence**

Collective intelligence, or web-enabled collective intelligence which is more pertinent to crowdsourcing, relates to the intelligence emerging from collaboration, collective efforts or competition in large, loosely organized groups of people (Cahalane et al. 2014; Malone et al. 2010). The process of harnessing collective intelligence includes several dimensions. The two fundamental types of things that crowds do are Create and Decide (Malone et al. 2010). For Create, collection (independent works assigned to the crowd), contest (only one or a few solutions for final outcome) and collaboration (interdependent tasks done by the crowd and managed by the firm) are the core strategies (called genes) (Malone et al. 2010). For Decide, group decision (require consensus within the group) and individual decisions are the core practices. We use this framework to understand and conceptualize our study context.¹

**Typology and Hypotheses**

**Experience Typology**

We first present a typology of crowd members (participants) in crowdsourced new product development. We define each type based on the distribution of prior experiences in crowdsourcing tasks. Although there are two generic types (i.e., generalists and specialists), the hybrid forms may emerge depending on the concentration of experiences with respect to some focal knowledge requirement (or task). Specifically, we use members’ prior experiences in other crowdsourced product development projects to quantify three metrics: diverse experience, specialized experience and concentration of experience.

Diverse experience is defined as the number of distinct tasks the member has experienced in past crowdsourced product development projects. Specialized experience is defined as the amount of prior experience the member has obtained in terms of the focal tasks. Here focal tasks refer to the tasks the member participates in for the current product development project. However, using only diverse

¹ For more details about the framework, see Malone et al. (2010).
experience and specialized experience cannot portray the whole picture regarding generalists and specialists. If one has both high diverse experience and high specialized experience, the domain in which the experiences are concentrated will matter. Concentrated experiences in limited domains renders one a T-shaped professional, while equally distributed specialized experience on all domains leads to an omniscient expert (Hansen and Von Oetinger 2001). Thus, in our study, concentration of experience is a measure of the extent of concentration of all the experiences the member has obtained in past product development tasks, which can be used to denote the balance between diverse and specialized experiences (Kang et al. 2012; Narayanan et al. 2009). Based on these three metrics, we may classify different types of crowd members in a crowdsourced new product development process. Table 1 shows the theoretical typology of crowd members according to combinations of values for these metrics.

When the extent of diverse experience, specialized experience and experience concentration are all high, such a member can be classified as T-shaped with respect to the focal task requirements (Hansen and Von Oetinger 2001; Kang et al. 2012; Narayanan et al. 2009). But when experience concentration is low, which means that the experiences are sparsely distributed in different types of task, the member will have highly specialized experiences in many of the task types, indicating an omniscient type. In addition, the T-shaped member could also be of the third type, which exhibits concentrated specialized experiences in tasks other than the focal task requirements. The fourth type indicates a type of generalist, which has broad experiences across different tasks but few experiences in each of the different types of tasks. The fifth type represents a specialist in the focal tasks, with specialized experiences concentrated in the focal areas. In contrast to this (i.e., the sixth type), when the specialized experience on focal tasks is low, this means the experiences are concentrated on other non-focal tasks, indicating a specialist in some other tasks. The remaining combinations are distinct from the others. In terms of the seventh type, on the one hand, when all the three metrics are zero, this would indicate new comers who have no experience before participating in the current product development task (i.e., first timers). On the other hand, when the three metrics are not zero but low, these members are most likely triers who have only initially participated in some tasks. Therefore, the seventh type corresponds to a group of inexperienced members. However, it is not possible to identify any specific type when only the specialized experience is high because such a combination is theoretically impossible. Thus, a total of seven types exist in the typology. In the empirical analysis, we classify the type of crowd members based on such a typology.

<table>
<thead>
<tr>
<th>No.</th>
<th>Diverse Experience</th>
<th>Specialized Experience</th>
<th>Experience Concentration</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>T-Shape in focal tasks</td>
</tr>
<tr>
<td>2</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Omniscient</td>
</tr>
<tr>
<td>3</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>T-Shape in other tasks</td>
</tr>
<tr>
<td>4</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Generalist</td>
</tr>
<tr>
<td>5</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Specialist in focal tasks</td>
</tr>
<tr>
<td>6</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Specialist in other tasks</td>
</tr>
<tr>
<td>7</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>New Comers/Trier</td>
</tr>
</tbody>
</table>

Note: The typology consists of seven types. The remaining possibility (i.e., low in diverse experience, high in specialized experience and low in experience concentration) is theoretically impossible and does not indicate any meaningful type.

**Hypotheses**

Consistent with the criteria of our typology, our hypotheses are also based on the three metrics of experience and the types that emerge from them. Specifically, we provide theoretical arguments about diverse experience, specialized experience and experience concentration in crowdsourced new product development, and then predict how each type of crowd member would impact the product development process. To develop the hypotheses about the crowd, we apply the mechanisms from the diversity literature on teams and groups, and discuss both individual level (crowd member) effects from the creativity perspective and group level (crowd) effects from the information processing perspective.

In crowdsourced new product development, a group of individuals form a crowd and collectively work on tasks that are designed to crowdsourc solutions or ideas from the community. Since crowd members need to search their own knowledge bases based on their prior experience in the relevant tasks, diverse experiences would provide members with richer knowledge components and more available ideas or solutions (Weisberg 1999). The availability of ideas is regarded as the basis for generating novel outputs.
(Taylor and Greve 2006). Therefore, diversity of experience should help individual to search and combine existing knowledge to generate creative and novel works (Hwang et al. 2014; Taylor and Greve 2006; Wulf and Schmidt 1997), resulting in performance increases. In addition, participants in the crowd do not always work independently. Besides collective works, they also collaboratively commit to a joint outcome (i.e., the product) with interdependencies among their tasks, ideas and solutions (Malone et al. 2010). Members in the crowd will interact with each other and communicate each other like in a virtual group or community. From this perspective, members with high experience diversity will be able to facilitate knowledge transfer through more effective group communication (Rulke and Galaskiewicz 2000). Other members that may lack task related information and experience would be able to receive assistance from these members with high experience diversity (Paulini et al. 2013). Also, the total level of diversity in terms of task experiences in the crowd can be further increased by experience diversity because of the interaction and broader exposition of knowledge (Nonaka and Takeuchi 1995), eventually leading to more creative ideas or solutions from the crowd. Thus, we argue that the experience diversity is positively associated with crowd performance (Huckman et al. 2009).

In terms of the crowd member types defined in our typology, four types in total are associated with high diverse experience. However, we notice that only the type “generalist” independently conforms to our argument for diverse experience, while others such as “T-shaped specialist” and “Omniscient” are also connected with the other two metrics. For generalists, they serve as the source of ideas and the driver of knowledge transfer to increase the diversity and creativity of the crowd (Hwang et al. 2014; Perry-Smith and Shalley 2003; Rulke and Galaskiewicz 2000). Thus, a crowd with more generalists will have better performance. Since our crowd composition is conceptualized as member types within the crowd, we define the baseline as new comers who do not have any experience in any dimension and use the proportion of member types to explain the effects. Therefore, for a given crowd size, a higher proportion of generalists will benefit the crowd in terms of performance. We propose:

H1: A crowd with higher proportion of generalists will have better performance.

Besides diverse experience, crowd members also rely on specialized experience to come up with ideas or solutions for the crowdsourced tasks. From the collective view, an individual in the crowd has to figure out the problem in the task and search from her existing knowledge to produce some output. Although specialized experience does not increase diversity, it could reduce the cost for the crowd member to work on the final output (Huang et al. 2012; Narayanan et al. 2009). Therefore, even though task specialization does not necessarily lead to novel or creative works, the efficiency and quality of the task output could be higher. Furthermore, although diversity can help to generate more ideas, it sometimes leads to unexpected outcomes due to the immaturity or uncertainty of ideas generated (Taylor and Greve 2006). Deep knowledge, from this view, can help individuals in the crowd effectively combine their knowledge and make their solutions more feasible (Hwang et al. 2014). Thus, the ideas or solutions generated by the individuals in the crowd may be better when they possess deep expertise in the tasks. In addition, the same expectation happens from the collaborative perspective (i.e., crowd level). Members with high level of specialized experience can provide practical suggestions and information about idea improvement for the works of other members who do not possess deep knowledge (Paulini et al. 2013). The deep knowledge shared by these members will help to make the ideas or solutions from the crowd more feasible and reliable, reducing the uncertainty caused by diversity and creativity in the product development (West 2002). Therefore, we expect that specialized experience is positively associated with crowd performance.

The two types of crowd members in our typology that conform to our arguments about specialized experience are “Specialists in the focal tasks” and “Specialists in non-focal (other) tasks”. Focal tasks, as defined earlier in our typology discussion, refer to the tasks the crowd member participated in the development of other products. Thus, having more experience in focal tasks will not only guarantee individual’s own work quality but also increase the reliability of others’ works more effectively. This effect, on the other hand, may not be salient for specialized experiences in non-focal tasks due to the absence of direct interaction and specialization. Following the same baseline in H1, we propose:

H2: A crowd with higher proportion of specialists in focal-tasks will have better performance.

In addition to diverse and specialized experiences that affect the crowd performance directly, the concentration of experience may also matter. The concentration of experience shifts the interaction between diversity and specialization, leading to more nuanced types of individuals in the crowd. When
both diversity and specialization are high, the cognitive load of individuals (i.e., too much knowledge) may be harmful for effective output and a balance of diversity and specialization (i.e., T-shaped) is suggested (Narayanan et al. 2009). Also, a concentrated focus on limited knowledge domains may enhance the learning outcome (Yang et al. 2008). However, given the innovative nature of tasks and voluntary participation in crowdsourced new product development, the balance between diversity and specialization may not be necessary. In a context of innovation, both diverse experience and specialized knowledge are important. Without deep knowledge in broad domains, individuals cannot combine their knowledge effectively to reduce the uncertainty of creative works and generate reliable solutions (Hwang et al. 2014; Novick 1988). Furthermore, due to the interdependencies among tasks in crowd-based product development (Malone et al. 2010), limited specialization in the focal domains will be insufficient because it is necessary to refer to other tasks or other crowd works to effectively search from own knowledge, which requires some understanding of non-focal tasks. Moreover, from the nature of voluntary participation, participants in the crowd could voluntarily choose whether or not to work on the tasks, so the effect of cognitive or learning load cannot be salient in this context (Archak and Ghose 2010; Huang et al. 2012). Thus, we argue that the concentration of experience will not matter in crowd performance.

The types of crowd members relevant to experience concentration are the “T-shaped” and “Omniscient” members. Combining our discussion about diverse experience and specialized experience, both of these types have the potential to enhance the crowd performance. Thus, crowd types with both high levels of diverse experience and specialized experience will lead to similar prediction. The two types – T-shaped specialist in focal and other tasks – possess both high levels of diversity and depth in prior experience. Similarly, the omniscient type members have high levels of specialization in most of the domains. Consistent with our discussion about specialists in non-focal tasks, the effect of specialization is not salient for T-shaped specialist in other tasks either, but diverse experience is expected to matter. Given our discussion about experience concentration, we do not expect significant differences due to concentration. Thus, in line with the same baseline condition in H1 and H2, we propose:

H3a: A crowd with higher proportion of T-shaped members in focal tasks will have better performance.
H3b: A crowd with higher proportion of T-shaped members in other tasks will have better performance.
H3c: A crowd with higher proportion of omniscient members will have better performance.

Data and Method

Study Context

Our empirical context is new product development at Quirky.com, a crowdsourcing platform for online inventions. As a community-driven new product development company, Quirky uses a crowdsourcing approach for both its product portfolio and product development. Users can submit their product ideas using on a problem-solution paradigm. Then Quirky with its community members will choose the promising ideas and start to develop the products. During the development process, Quirky also crowdsources important product development related tasks and decisions from community members through different development projects. When a product is released into the real world, Quirky will sell the product through various channels (e.g., via partnerships with major brick and mortar retailers such as Target, Best Buy and Bed Bath & Beyond as well as through Quirky’s own e-commerce website). Anyone who has contributed to the development of the product, including the initial inventor (i.e., the individual who submitted the original product idea), community members (i.e., those who participated in the product development projects), and Quirky (who is in charge of manufacturing) will share the proceeds from the product sales. The most successful crowd member has earned more than USD $800,000 at Quirky. Quirky’s business model has attracted a lot of venture capitalists and participation from numerous inventors in the community (more than 1,000,000 members) (Colao 2013).

The first stage of business model, called Invent, comprises the idea generation process where inventions from ordinary people are submitted. It includes both competition among ideators and collaboration from the community through comments, identification of similar products and social interactions among

---

members. If an idea is selected for development, it moves into the Influence stage, where Quirky sets up a series of collaborative projects based on different tasks that are crowdsourced to the community. When a product successfully completes its development (on paper or as a prototype), it will be launched into Concept Portfolio. Our study focus is the second stage (i.e., the Influence stage), which comprises the crowdsourced product development process. The process Quirky adopts for harnessing collective intelligence is Collection and Collaboration (Malone et al. 2010). At the project level, Quirky crowdsources a specific task to the community and community members independently work on the solutions, indicating the Collection mode. But at the product level, different tasks are interdependent and members should consider others’ works in order to complete their own work, consistent with the Collaboration mode. Although Quirky emphasizes the concept of team in the Influence stage, we find that it may not completely fit the term “team” (Dissanayake et al. 2014) because of the mixture of modes. Instead, we use “crowd” to characterize the participants in a whole product development process, which is consistent with the call for the investigation of the “crowd” (Pedersen et al. 2013).

Data Collection

We retrieved all product information and product development information using a software crawler. The details of crowdsourced projects within each product development process were also collected, including the type, task, submissions and comments for each project. Finally, user profiles and information related to the idea of each product were collected to build a holistic picture of the product development process. The time window of our data is from May 2009, the start of Quirky’s operations, to June 2014. The full dataset includes 828 products, with 3,044 projects for these products, as well as 33,789 unique users who participated in the crowdsourced product development process. In our analysis, we focus only on those products with community participations (569 products) so that a crowd is constructed and used in the product development process. These were used for the measurement of members’ prior experiences, which are then used to identify the types of members in terms of experience. Furthermore, some products were not finished (i.e., not launched into the Concept Portfolio) due to some unexpected development difficulties or due to the censored time window of our data collection. Since we do not have complete information about these products, they were not included in the analysis. Nonetheless crowd members’ experiences in these products were included. We also only use data starting from 2010 since members on the platform generally do not have any prior experiences in the first year.3 Finally, we excluded repeated products which are based on the same idea since they usually have duplicated development processes. Outliers and other incomplete information records in the data were also examined and excluded. Our final sample includes 413 products that were developed at least on paper between 2010 and 2014.

Measures

To measure the prior experience, we first identified the crowd members in each product’s development. Specifically, we use submissions as the criterion for member selection. A user is considered a crowd member if he or she submits something in the crowdsourced projects during product development. Although peripheral contributions exist in the product development, they are usually not creative works and not influential. Future work may examine the significance of peripheral contributions in this context. Thus, we set the criterion for crowd and measure experiences based on user contributions within each product development. This procedure leaves us with 274,281 observations of product-user pairs.

We measure members’ past experience based on the three metrics – diverse experience, specialized experience and concentration. Similar to related studies (Pedersen et al. 2013), we note that the crowd in the product development crowdsourcing is dynamically constructed. Also, the experience of members is also dynamic and product-crowd specific. Thus, we computed the each member’s experience in each product using the join time, which is the time at which the member started to participate in the product development, as the cutoff time for experience. In addition, experience is at the product level, which means that all relevant experiences for a product development process can be included in a member’s experience portfolio only after the product development completes (or is stopped prematurely; or is censored due to the time window of data collection). This is because in the product development context,

---

3 That said, data from 2009 were used in capturing the experiences of crowd members that participated in products from 2010 and onwards.
the product is a whole entity or unit for acquiring complete experiences (i.e., partial experiences are not counted). This procedure allows us to capture both the dynamic crowd formation process as well as differences in experience accumulation for each user within a particular crowd.

In our context, there are five different types of projects corresponding to five specific tasks (i.e., a 1-to-1 mapping). Tasks include Research, Design, Styling, Naming and Tagline setting. Thus, we operationalized diverse experience as the number of distinct tasks the user has participated in past product development projects. Specialized experience was operationalized as the amount of prior experiences in the focal tasks. Specifically, it could be written as \( \text{SpecializedExperience}_i = \frac{1}{P} \sum_j \text{Exp}_{ij} \), where \( P \) is the number of task types crowd member \( i \) has participated in for the focal product development, \( T_i \) is the total number of products development in which the user has participated, and \( \text{Exp}_{ij} \) is the number of focal tasks the user participated in the development of product \( j \).

To measure the concentration of experience of crowd members, we use the Herfindahl-Hirschman Experience Index (HHEI), which is derived from the Herfindahl-Hirschman Index frequently used to measure market concentration in economic studies. This measurement has been used in software development studies to represent the concentration of experience (Kang et al. 2012; Narayanan et al. 2009). Specifically, the computation of experience concentration by HHEI is \( HHEI = \frac{1}{N} \sum_{i=1}^{N} \frac{\text{Exp}_i}{\text{SumExp}} \) , where \( N \) is the total number of distinct tasks (\( N=5 \) in our context) in all products’ development, \( \text{Exp}_i \) is the number of task experiences of member \( i \) in task type \( k \), and \( \text{SumExp}_i \) is number of all tasks crowd member \( i \) participated in the past across all the task areas.

**Identifying Experience-based Crowd Member Types**

Using the operationalizations of the three metrics, we computed the time-varying experience measures for each crowd member within each product development process. Then a clustering approach was used to identify the types of members for each product-crowd. Cluster analysis has been widely used in related literature to empirically capture types of users given patterns of observable behaviors (Hahn and Lee 2013; Lin et al. 2014). We adopt the two-step procedure in Ketchen and Shook (1996) to conduct the cluster analysis. First, we performed hierarchical clustering (Maimon and Rokach 2005) and investigate the dendrogram to determine a suitable number for clusters. We also validate the solution using the \( k\)-means clustering approach. Specifically, we plot the number of clusters against the within-cluster sum of squared errors, and verify the number of clusters based on the “kink” and stability of the curve (Ketchen and Shook 1996; Lin et al. 2014). After deciding the number of clusters, we use the \( k\)-means approach to assign the membership of observations to clusters on the 274,281 product-member pairs to capture the dynamic patterns of experiences. Finally, we matched the clustering results to our typology.

**Empirical Model**

We use regression analysis to test the impact of each type of crowd members on their product development performance. We specify the variables used in our empirical model as follows.

**Dependent Variable**

Development Duration (\( \text{Duration} \)): We measure product development performance based on the duration of the crowdsourced product development projects.\(^4\) Several reasons support the use development project duration to measure crowd performance. First, in a project with specific tasks, Quirky (usually a team assigned for the product) will review the submissions and close the project when it could utilize the solutions for development decisions. Therefore, the duration captures how Quirky perceives the quality the crowdsourced work. Second, in the product development process, Quirky will typically control the quality of product and launch the product into the Concept Portfolio stage only if reaches an acceptable level of quality. Thus, all else being equal, a shorter duration means that Quirky spends less time on

---

\(^4\) Ideally, the dependent variable should be the overall productivity of the crowd. Unfortunately, it is not possible to observe the crowd members’ actual efforts expended to measure the internal crowd-level productivity with only publicly observable data on product development.
examining the crowdsourced solutions and submissions from multiple tasks, implying better performance from the crowd. In contrast, a longer development duration implies that the crowd did not generate satisfactory solutions for Quirky and as a result took Quirky more time for overall product development (i.e., close the projects and launch the product). However, since the crowd dynamically emerges, using the product development duration from its start to end may raise some confounding and causality issues (i.e., reverse causality whereby a longer time causes more participants and more experiences of participants). To address this issue, we adopt an aggregation approach for the operationalization. Since the crowdsourced product development (i.e., crowd work) is comprised of several development projects and solutions are typically submitted during the initial days in these projects, the durations of all development projects are used. Specifically, we computed the duration of each of the projects and summed them to derive the duration at the product level. Such an operationalization helps to avoid the issues caused by the dynamic nature of the crowd and does not affect the validity of the crowd performance measure.

Independent Variables

Type of crowd member: To examine the impact of experience types on crowd performance, we used the proportion of each experience-based user type in the crowd as the independent variables in our regression models (Hahn and Lee 2013; Inbar and Barzilay 2014). The types are determined by the empirical identification from the cluster analysis.

Control Variables

Given that our dependent variable for crowd performance is product development duration, we need to control for several factors that not only affect crowd performance but also product development time.

Number of Submissions (Submissions): The number of submissions is the number of solutions, ideas and other creative works submitted by crowd members in all of the projects for each product development. In our context, because each member can only submit a limited number of solutions in each project, the number of submissions cannot measure the performance of crowd. Therefore, we include it as a control for the amount of crowd work. In addition, the number of submissions also accounts for the size of the crowd because each user can only submit a limited number of solutions.

Number of Projects (NumProjects): Since we use the sum of project durations as the dependent variable, it is necessary to control for the number of projects. In addition, the number of projects indicates the amount of required crowdsourced work for product development, which is also useful for explaining the duration of product development.

Brainstorm (HasBrainstorm): Brainstorming is an offline activity for product development using an expert panel. Typically, if a brainstorm session is conducted, a section with a video is displayed in the product timeline. Since brainstorming is done before the start of product development, we can use it as a control variable. We use an indicator variable for whether the product used a brainstorming session.

Inventor Products (InventorProducts): We also control the characteristics of the product inventor (idea submitter). Specifically, we use the number of prior successful ideas (moving to development) as a control.

Number of Comments (Comments): We control the total number of comments on submissions by crowd members during the product development process for the interaction and collaboration of the crowds.

Average Crowd Ideation Influence (AvgIdeaInfluence): Since each crowd could be different due to members’ own intelligence levels and other experiences, it is necessary to account for the confounding effect caused by the crowd intelligence. Specifically, we utilize the performance of crowd members in the Invent stage to account for this. We compute the cumulative amount of influence points (a measure of contribution for ideation defined by Quirky) earned by each crowd member before he/she joined the crowd, and then take the average across the members within a crowd.

Product-Specific Factors: The product development process is usually affected by the complexity and uniqueness of the product. Since it is difficult to know a-priori the complexity of the final product, we use three proxy variables to capture product-level heterogeneity. We use the number of comments, number of similar products and length of solution in idea description to control for product characteristics. First, we control for the number of comments in the ideation (or invent) stage of the product (IdeaComments). We only compute the number of comments before the product moves into development (i.e., prior to selection
for product development). More community feedback may indicate that the product has more elements to be discussed, reflecting greater complexity. Next, we control for the number of similar products submitted by community members about the product (SimilarProducts). If more similar products exist, the uniqueness of the product idea is lower but the number of elements of the product may be higher. We use this to control for the uniqueness of product. Finally, the length of solution provided by the ideators is controlled for the complexity of product idea (Solution). Since idea submitters follow a problem-solution paradigm, the solution indicates the possible product designs and signals the complexity of the problem.

Category and Time: We also control the category of each product using dummy variables. We classify the eight categories into three main categories: electronic-related, home-related and play-related. Also, year dummies are used to control for time trends, policy change and year-specific effects.

In the empirical analysis, we log-transform product development duration, number of submissions, number of comments, average crowd ideation influence, number of ideation comments, number of similar products and length of solution because of their range and skewed distributions. Since our data is cross-sectional at the product-crowd level, we used OLS to estimate the parameters in the regression model.

Results

Clustering Results

We first performed hierarchical clustering to determine the suitable number of clusters. However, given that our sample includes 274,281 observations, it is not practical to employ hierarchical clustering due to computational limitation. To overcome it, we employed a bootstrapping approach where we selected a random sample of 10,000 observations to perform hierarchical clustering and repeat the process with different random subsamples. The dendrograms indicate that a six clusters solution is stable across bootstrapped samples. We then performed k-means cluster analysis on the full sample using six clusters solution. To further verify the results, we also plotted the within cluster sum of squared errors against the number of clusters (see Figure 1). The curve stabilizes at the k=6 point, and a significant “kink” is clearly observable there. This further supports our six clusters solution (Thorndike 1953). Finally, as k-means approach randomly selects the initial values of cluster centroid (starting points), we ran the analysis for the six cluster solution using multiple random starting points (1,000 replications) and the clusters are quite stable in terms of cluster centroid and size (reliability > 0.8).

![Figure 1: Plot of Within Cluster Sum of Squared Errors against Number of Clusters](image)

5 The hierarchical clustering requires the distance matrix between each observation, which will generate $N(N-1)/2$ pairs in the matrix. Given that our sample size (274,281), it is impossible to address the exponentially increased number of pairs in the matrix. Thus, we follow the existing literature to use random subsamples. In addition, due to the dynamic nature in our sample, we use k-means approach on the whole sample to verify the clusters.
Table 2 summarizes the identified clusters and the corresponding crowd member types. We code each cluster by comparing the mean of each metric within cluster with their mean for the whole sample to match our typology. Interestingly, not all types in the theoretical typology were identified from our data. New comers (Cluster4)/Triers (Cluster6), T-Shape in other tasks (Cluster1) and Generalist (Cluster3) are matched to the types in our typology. However, we found two types of omniscient members (i.e., those who have specialized experience in most task areas) in our data – Cluster 2 and Cluster 5. Both clusters have high levels of experience and low levels of experience concentration. Cluster2 shows a group of crowd members with high level of both diverse experience and specialized experience, but a moderate level of experience compared to Cluster5. We code Cluster2 as “Deep Generalist” (Hwang et al. 2014) and Cluster5 as “Omniscient” (Kang et al. 2012) because they are statistically separable (SpecializedExperience: t=257.43, p<0.0001). In addition, we note that Cluster6 could be specialist in other tasks (type 6 in the typology) based on the three experience metrics. But when considering the total product level experiences of this cluster, we note that their experiences are quite limited (around 1.95). Thus, they are actually more representative of triers instead of specialist in other tasks. The reason for their high experience concentration is that the computation of HHEI would cause extremely high value when a member has only one or two experience in one or two task areas (i.e., one unit of experience in one area does not necessarily mean a high concentration, even though the HHEI computation results in an extremely high value). To further verify this, we computed a balanced measure of HHEI by adding 1 experience in each task area before computing HHEI. The clustering results can distinguish the cluster of triers with low experience concentration from specialists in other tasks. Thus, a total of six types of members were identified based on experience breadth, depth and concentration from the empirical data.

### Table 2: Clustering Results and Corresponding Types

<table>
<thead>
<tr>
<th>Type</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
<th>Cluster 4</th>
<th>Cluster 5</th>
<th>Cluster 6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-shaped in non-focal</td>
<td>62615</td>
<td>25100</td>
<td>73100</td>
<td>96130</td>
<td>6652</td>
<td>15184</td>
<td>274281</td>
</tr>
<tr>
<td>Deep generalist</td>
<td>3.7656</td>
<td>4.9786</td>
<td>5.0000</td>
<td>4.9794</td>
<td>1.6576</td>
<td>2.8604</td>
<td></td>
</tr>
<tr>
<td>Generalist</td>
<td>8.9065</td>
<td>99.0516</td>
<td>23.3769</td>
<td>0.0000</td>
<td>220.5670</td>
<td>1.0920</td>
<td>22.7500</td>
</tr>
<tr>
<td>New comers</td>
<td>0.3087</td>
<td>0.2553</td>
<td>0.2556</td>
<td>0.2534</td>
<td>0.6872</td>
<td>0.2730</td>
<td></td>
</tr>
<tr>
<td>Omniscient</td>
<td>0.3087</td>
<td>12.7359</td>
<td>36.6716</td>
<td>0.0000</td>
<td>277.8890</td>
<td>1.9472</td>
<td>31.9702</td>
</tr>
<tr>
<td>Tiers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Pair-wise comparison between each cluster using their mean and standard deviation are conducted. Except 2-5 for diverse experience and 2-3 for experience concentration, all of the pairwise differences are significant at p<0.01 level.

### Analysis Results

Using the results of the cluster analysis, the independent variables were operationalized as the proportion of the five types (out of the six identified). Thus, the proportion of T-shaped in other tasks members (TriOther), deep generalists (DeepGeneralist), generalist members (Generalist), Omniscient members (Omniscient) and Triers (Triers) in the crowd are the five independent variables and the largest type, the new comer type, is used as the base case benchmark. Table 3 shows the descriptive statistics of key variables. We estimated our parameters progressively by first estimating a model with control variables only (Model 1) and then adding the independent variables of interest (i.e., crowd type variables) in Model 2 (main analysis). We also conducted robustness checks to verify our results from Model 3 to Model 5. Heteroskedasticity-robust standard errors are used in the estimation. Variance Inflation Factors (VIFs) were checked and below the recommended thresholds (Belsley et al. 2005; Cohen et al. 2013). The regression results are presented in Table 4.

In Model 1, we only include the control variables. We observe that the effect of number of submissions is negative and significant (Submissions: $\beta=-0.409$, $p<0.01$), which means larger crowd and more crowd works facilitate (i.e., shorten) crowdsourced product development. Attracting more crowd participants and submissions help to gather more solutions and indicate better crowd performance. As an important component of development duration, it is not surprising that the coefficient of NumProjects is positive and significant ($\beta=0.159$, $p<0.01$). The coefficient of HasBrainstorm is negative but not significant ($\beta=-0.112$, ns). Although the brainstorm section could help crowd understand the product and be an indicator of product feasibility, it does not seem to significantly affect crowd performance in product development. The coefficient of ProductInventor is not significant ($\beta=-0.00543$, ns), showing no effect of inventor’s...
past successes. Interestingly, the number of comments in the development projects is positively associated with the product development time (Comments: $\beta=0.139$, $p<0.01$), suggesting that more discussions among the crowd members through comments does not lead to a better performance. One possibility is that the number of comments reflects the level of consensus in collaboration – more comments indicate that it is difficult for crowd participants to achieve consensus in the collaboration. Further examination of this variable should be conducted in future research. The coefficient of AvgIdeaInfluence in Model 1 is negatively significant ($\beta=-0.585$, $p<0.01$), showing that averagely more intelligent crowd would perform better by reducing development time. In terms of the three proxy variables for product heterogeneity, they are for the most part positively associated with product development time (IdeaComments: $\beta=0.0729$, $p<0.05$; SimilarProducts: $\beta=0.388$, $p<0.01$; Solution: $\beta=0.0310$, ns). More feedbacks through comments during the ideation stage indicates more discussions and suggestions, which could be related to product complexity, while more similar product identified may be an indicator of product uniqueness, which is also relevant to product complexity. The length of solution is not significant, which means the solutions provided by the ideators do not affect crowd performance in product development. However, there are still potential competing effects from these three variables even though some of the coefficients were found to be positive and significant. Additional investigations on the content of comments, similar products and idea description should be conducted in future research.

In Model 2, we add the crowd member type variables. The control variables are generally stable except for IdeaComments and AvgIdeaInfluence. The insignificance of AvgIdeaInfluence may imply that when controlling for the crowd members’ experience and composition, the crowd intelligence level and other experience does not matter much, highlighting the uniqueness of the research context. In Model 2, we first observe that the type that contributes most to product development performance is the omniscient one. The coefficient of Omniscient is negatively significant ($\beta=-6.597$, $p<0.05$). This is not that surprising since members with high level of diverse experiences and specialized experiences should be more knowledgeable than others. Thus, a crowd with a higher proportion of such members would outperform others, which is consistent with the supportive evidence of diverse experiences and specialized experiences. Consistently, the deep generalists also have positive effects on crowd product development performance (DeepGeneralist: $\beta=-2.645$, $p<0.05$). But the magnitude of this type is smaller than the omniscient type. These two results suggest that H3c is supported. We also find that compared to the new comers, the type of T-shaped in other tasks can also facilitate product development process, although not as much as the deep generalist members (TinOther: $\beta=-1.308$, $p<0.1$). Members of this type have relatively broad experience but do not have deep experience in the focal tasks. However, they also have some experience in other task areas. Such members may have stronger motivations to explore their experiences in the focal tasks and their experience in other tasks would help them in the focal task areas (Amabile 1983; Taylor and Greve 2006). So it seems that in crowdsourced product development, crowdsourcers should not only attract highly experienced participants but also those who are T-shaped in other tasks. This could be achieved by increasing the relatedness of tasks across different task areas. H3b is supported by this finding.

**Table 3: Descriptive Statistics**

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td>413</td>
<td>155.3</td>
<td>164.6</td>
<td>10</td>
<td>1,024</td>
</tr>
<tr>
<td>TinOther</td>
<td>413</td>
<td>0.235</td>
<td>0.0606</td>
<td>0.0800</td>
<td>0.426</td>
</tr>
<tr>
<td>DeepGeneralist</td>
<td>413</td>
<td>0.0828</td>
<td>0.0579</td>
<td>0</td>
<td>0.331</td>
</tr>
<tr>
<td>Generalist</td>
<td>413</td>
<td>0.292</td>
<td>0.0844</td>
<td>0.128</td>
<td>0.720</td>
</tr>
<tr>
<td>Omniscient</td>
<td>413</td>
<td>0.0157</td>
<td>0.0227</td>
<td>0</td>
<td>0.0988</td>
</tr>
<tr>
<td>Trier</td>
<td>413</td>
<td>0.0478</td>
<td>0.0262</td>
<td>0</td>
<td>0.164</td>
</tr>
<tr>
<td>Submissions</td>
<td>413</td>
<td>1.879</td>
<td>1.483</td>
<td>38</td>
<td>8,041</td>
</tr>
<tr>
<td>NumProjects</td>
<td>413</td>
<td>4.775</td>
<td>1.400</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>HasBrainstorm</td>
<td>413</td>
<td>0.121</td>
<td>0.327</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>InventorProducts</td>
<td>413</td>
<td>0.627</td>
<td>1.667</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Comments</td>
<td>413</td>
<td>40.98</td>
<td>39.63</td>
<td>0</td>
<td>260</td>
</tr>
<tr>
<td>AvgIdeaInfluence</td>
<td>413</td>
<td>4.961</td>
<td>2.474</td>
<td>0.865</td>
<td>19.51</td>
</tr>
<tr>
<td>IdeaComments</td>
<td>413</td>
<td>344.1</td>
<td>295.9</td>
<td>0</td>
<td>1,906</td>
</tr>
<tr>
<td>SimilarProducts</td>
<td>413</td>
<td>6.247</td>
<td>4.167</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Solution</td>
<td>413</td>
<td>192.1</td>
<td>82.51</td>
<td>0</td>
<td>244</td>
</tr>
</tbody>
</table>

In Model 2, we add the crowd member type variables. The control variables are generally stable except for IdeaComments and AvgIdeaInfluence. The insignificance of AvgIdeaInfluence may imply that when controlling for the crowd members’ experience and composition, the crowd intelligence level and other experience does not matter much, highlighting the uniqueness of the research context. In Model 2, we first observe that the type that contributes most to product development performance is the omniscient one. The coefficient of Omniscient is negatively significant ($\beta=-6.597$, $p<0.05$). This is not that surprising since members with high level of diverse experiences and specialized experiences should be more knowledgeable than others. Thus, a crowd with a higher proportion of such members would outperform others, which is consistent with the supportive evidence of diverse experiences and specialized experiences. Consistently, the deep generalists also have positive effects on crowd product development performance (DeepGeneralist: $\beta=-2.645$, $p<0.05$). But the magnitude of this type is smaller than the omniscient type. These two results suggest that H3c is supported. We also find that compared to the new comers, the type of T-shaped in other tasks can also facilitate product development process, although not as much as the deep generalist members (TinOther: $\beta=-1.308$, $p<0.1$). Members of this type have relatively broad experience but do not have deep experience in the focal tasks. However, they also have some experience in other task areas. Such members may have stronger motivations to explore their experiences in the focal tasks and their experience in other tasks would help them in the focal task areas (Amabile 1983; Taylor and Greve 2006). So it seems that in crowdsourced product development, crowdsourcers should not only attract highly experienced participants but also those who are T-shaped in other tasks. This could be achieved by increasing the relatedness of tasks across different task areas. H3b is supported by this finding.
Crowd Experience in Crowdsourced New Product Development

Table 4: Regression Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>TnOther</td>
<td>-1.308***</td>
<td>-1.627***</td>
<td>-1.235**</td>
<td>-1.401**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.689)</td>
<td>(0.660)</td>
<td>(0.683)</td>
<td>(0.695)</td>
<td></td>
</tr>
<tr>
<td>DeepGeneralist</td>
<td>-2.645**</td>
<td>-3.222**</td>
<td>-2.563**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.328)</td>
<td>(1.345)</td>
<td>(1.315)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalist</td>
<td>-0.254**</td>
<td>0.0305</td>
<td>-0.283</td>
<td>-0.314</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.840)</td>
<td>(0.882)</td>
<td>(0.825)</td>
<td>(0.826)</td>
<td></td>
</tr>
<tr>
<td>Omniscient</td>
<td>-6.597**</td>
<td>-7.155**</td>
<td>-7.438**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.913)</td>
<td>(2.871)</td>
<td>(2.885)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trier</td>
<td>-0.393</td>
<td>-0.545</td>
<td>-0.418</td>
<td></td>
<td>-3.654***</td>
</tr>
<tr>
<td></td>
<td>(1.604)</td>
<td>(1.606)</td>
<td></td>
<td>(1.022)</td>
<td></td>
</tr>
<tr>
<td>Omniscient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-3.654***</td>
</tr>
<tr>
<td>ln(Submission)</td>
<td>-0.409***</td>
<td>-0.224**</td>
<td>-0.217**</td>
<td>-0.231**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0640)</td>
<td>(0.106)</td>
<td>(0.105)</td>
<td>(0.104)</td>
<td></td>
</tr>
<tr>
<td>ln(Members)</td>
<td></td>
<td>-0.186</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.141)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NumProjects</td>
<td>0.159***</td>
<td>0.0996***</td>
<td>0.0803**</td>
<td>0.0923**</td>
<td>0.108***</td>
</tr>
<tr>
<td></td>
<td>(0.0356)</td>
<td>(0.0379)</td>
<td>(0.0375)</td>
<td>(0.0384)</td>
<td>(0.0376)</td>
</tr>
<tr>
<td>HasBrainstorm</td>
<td>-0.112</td>
<td>-0.102</td>
<td>-0.107</td>
<td>-0.102</td>
<td>-0.119</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.104)</td>
<td>(0.105)</td>
<td>(0.104)</td>
<td>(0.104)</td>
</tr>
<tr>
<td>InventorProducts</td>
<td>-0.00543</td>
<td>-0.00776</td>
<td>-0.00862</td>
<td>-0.00525</td>
<td>-0.00885</td>
</tr>
<tr>
<td></td>
<td>(0.0262)</td>
<td>(0.0258)</td>
<td>(0.0263)</td>
<td>(0.0256)</td>
<td>(0.0258)</td>
</tr>
<tr>
<td>ln(Comments)</td>
<td>0.139***</td>
<td>0.135**</td>
<td>0.141**</td>
<td>0.137***</td>
<td>0.137***</td>
</tr>
<tr>
<td></td>
<td>(0.0419)</td>
<td>(0.0380)</td>
<td>(0.0384)</td>
<td>(0.0371)</td>
<td>(0.0386)</td>
</tr>
<tr>
<td>ln(AvgIdeaInfluence)</td>
<td>-0.585***</td>
<td>-0.226</td>
<td>-0.173</td>
<td>-0.225</td>
<td>-0.444</td>
</tr>
<tr>
<td></td>
<td>(0.139)</td>
<td>(0.192)</td>
<td>(0.199)</td>
<td>(0.190)</td>
<td>(0.184)</td>
</tr>
<tr>
<td>ln(IdeaComments)</td>
<td>0.0729**</td>
<td>0.0527</td>
<td>0.0544</td>
<td>0.0448</td>
<td>0.0517</td>
</tr>
<tr>
<td></td>
<td>(0.0338)</td>
<td>(0.0335)</td>
<td>(0.0336)</td>
<td>(0.0342)</td>
<td>(0.0337)</td>
</tr>
<tr>
<td>ln(SimilarProducts)</td>
<td>0.388***</td>
<td>0.427**</td>
<td>0.432**</td>
<td>0.438***</td>
<td>0.433**</td>
</tr>
<tr>
<td></td>
<td>(0.0712)</td>
<td>(0.0782)</td>
<td>(0.0807)</td>
<td>(0.0776)</td>
<td>(0.0772)</td>
</tr>
<tr>
<td>ln(Solution)</td>
<td>0.0310</td>
<td>0.0253</td>
<td>0.0278</td>
<td>0.0109</td>
<td>0.0215</td>
</tr>
<tr>
<td></td>
<td>(0.0251)</td>
<td>(0.0254)</td>
<td>(0.0258)</td>
<td>(0.0263)</td>
<td>(0.0243)</td>
</tr>
<tr>
<td>HasSale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.218**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.107)</td>
</tr>
<tr>
<td>Constant</td>
<td>5.962***</td>
<td>5.083***</td>
<td>4.582***</td>
<td>5.162***</td>
<td>4.980***</td>
</tr>
<tr>
<td></td>
<td>(0.633)</td>
<td>(0.945)</td>
<td>(1.077)</td>
<td>(0.939)</td>
<td>(0.936)</td>
</tr>
<tr>
<td>Observations</td>
<td>413</td>
<td>413</td>
<td>413</td>
<td>413</td>
<td>413</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.451</td>
<td>0.476</td>
<td>0.471</td>
<td>0.482</td>
<td>0.474</td>
</tr>
</tbody>
</table>

**Significance levels:** ***p<0.01, **p<0.05, *p<0.1

**Notes:** Category and time dummies are included. Robust standard errors in parentheses.

Interestingly, we do not find any other significant types in our main analysis. More generalists in the crowd do not affect product development performance (Generalist: β=-0.254, ns), which is not entirely consistent with the existing literature on generalist in group context. H1 is thus not supported. One possible reason may be the limited nature of communication in the crowd rooted in virtual community. Generalists are able to facilitate knowledge transfer and sharing within a group, but this benefit would only materialize if there is vivid communication among the members, which is not the case in these crowds. In addition, in our study context, the number of task areas is limited such that the opportunities for learning new knowledge of generalists are constrained and as a result, they do not clearly outperform the new comers. On the other hand, new comers and triers may have more incentives in exploring the tasks than generalists who have already experienced all of the task areas, while generalists need more time to explore and develop deep expertise and evolve into other (more experienced) types over time (Cahalane et al. 2014). The coefficient of Trier is also not significant (β=-0.393, ns). Compared with new comers, although triers have a little more experiences in some tasks, their experiences are still limited and cannot
provide better submissions. However, H2 and H3a could not be directly tested from the results. H2 is only partially supported by our findings since our results suggest a positive effect of specialized experience. However, whether the effect is contingent on the diverse experience is unknown. H3a is also partially supported because our findings do not suggest performance benefits from concentrated experience in the crowd composition. Further evidence of these two hypotheses will be examined in our future work.

We noticed that the crowd type of cluster 1 (T-shaped in non-focal tasks), compared with type by cluster 2 (Deep generalist) and cluster 4 (Omniscient), do not have that much diversity in experience (Mean=3.77). This may indicate a moderate level of diverse experience outside our typology. However, given that in our context, only a total of five tasks have been defined, task diversity of 3.5 to 4 is actually quite high, especially when compared with new comers and triers. Nevertheless, it is still worth noting that this type of crowd member actually outperforms generalists, which contradicts to our expectations. Some initial analyses were conducted in the robustness check section but we will investigate this group in future work.

A number of additional analyses are conducted to check the robustness of the results. First, we use the number of members (Members) to formally control for the crowd size. Given that we use number of submissions (Submissions) in prior models and the two variables are highly correlated ($\rho=0.988$, $p<0.0001$), Model 3 indicates that the results are generally similar. This means the use of number of submissions to account for the crowd size is valid. Second, we observe that in our sample, some products have been promoted to the market while others have not. It may be the case if products are qualitatively different according to the final outcome. However, since this may have reverse causality issues (whether a product being in the market is affected by the development duration and crowd performance, or vice versa), we did not use it in the main analysis but include it as a post-hoc robustness check. By adding the variable HasSale to indicate whether a product has been in the market or not, Model 4 shows that the results are consistent and not affect by the qualitative difference of market promotion. Third, in order to better match our theoretical typology, we also tried a four-cluster classification that groups similar types (triers and new comers into inexperienced group, and deep generalist and omniscient ones into omniscient group). The results in Model 5 provide similar conclusions in terms of generalists, T-shaped in other tasks and omniscient ones. Finally, we also conducted additional initial analysis to identify the evolution of crowd experience. We find that participants in Cluster1 (T-Shaped in non-focal tasks) usually participate less than generalists (Cluster3) (on average 7 vs. 25 across product development; 2.45 vs. 2.55 within product development, $p<0.0001$). Although generalists are more active, they may be distracted by too many projects without having enough deep experiences. But for members in Cluster1, although they are not as active, they carefully explore their knowledge and perform relatively better. Additional analyses about the community is left for future research.

**Discussion and Conclusion**

In this study, we empirically investigated the type of crowd members in a new product development context. We developed a typology and corresponding hypotheses for the types of crowd members based on the diversity literature and the theoretical framework of generalists vs. specialists. With data on new product development at Quirky.com, we empirically identified six types of crowd members. Our empirical results showed that in addition to the omniscient members and deep generalists, members with T-shaped experiences in other task areas also benefit the crowd performance in terms of reducing the crowdsourced product development duration. Interestingly, generalists were found not to be very valuable.

Our study contributes in several aspects to theory and practice. First, this study contributes to the literature on crowdsourcing and open innovation. A large body of the crowdsourcing literature focuses primarily on contests, examining the motivation, antecedents and consequences (Boudreau et al. 2011; Hou et al. 2011; Yang et al. 2010; Yang et al. 2009). However, as an emergent component of crowdsourcing, the value of collaboration and co-creation are still understudied (Pedersen et al. 2013). The new forms of crowdsourcing have led to new business models (Song et al. 2009). Our study examines such a crowdsourcing model and empirically identifies the important drivers of success in crowdsourced new product development. The collaborative crowdsourcing process makes use of the interdependence among tasks and participants to further harness collective intelligence. The concept of crowd co-creation and collective intelligence have attracted much attentions in both crowdsourcing and crowdfunding contexts (Avital et al. 2014; Nickerson et al. 2014). Our study serves as one of the initial empirical works examining the key components and unpacks the crowd co-creation process in online communities.
Second, our study also contributes to the literature on virtual communities. We identified the important types of crowd members in a crowdsourcing community where community members work in both collective and collaborative manner. We find that generalists do not improve crowd co-creation works in a large virtual group (crowd) context possibly due to limited information transfer and communication. Members with T-shaped experience in other tasks actually were found to facilitate the product co-design or co-creation process through crowdsourcing. These findings provide further insights related to the impact of virtual community on value co-creation (Antorini et al. 2012). In addition, our work also captures the dynamics of crowd formation in crowdsourcing community (Pedersen et al. 2013) and the dynamic experience of crowd participants. More attention on dynamically participative crowds in virtual communities and their activities should be taken in future works.

Third, we contribute to the theory on generalists and specialists, and extend the diversity literature to large and dispersed online groups. In traditional group work such as software development, T-shaped experience distributions with specialized experiences in the focal knowledge domain is typically preferred (Kang et al. 2012; Narayanan et al. 2009). Furthermore, groups composed of generalists was also shown to exhibit better performance (Rulke and Galaskiewicz 2000). However, in our context, generalists were not found to be a beneficial type. When a generalist reaches the limit of knowledge areas, she may have a weaker motivation to improve performance and needs more time to further explore the depth of her knowledge portfolio to evolve into a more experienced type. Limited communication may also lead to unexpected implications. In the diversity literature, communication and coordination cost are discussed for large groups but there is a lack of empirical examination (Taylor and Greve 2006). We extend the theory about diversity in large and loose online groups and find that diversity may only have limited influence for innovation and creativity in the absence of knowledge depth. In addition, a T-shaped member in other tasks is also shown to be an important type in crowdsourced groups. Specialized experience may have the potential to transfer from other areas to the focal area when diverse experience is sufficient for such a transfer. Finally, our results show that only diverse experience may not work in our context. Only with specialized experience in various task areas could members collectively produce better performance, which is consistent with Hwang et al. (2014) in the context of individual idea innovation. The mechanisms of knowledge exploration and experience shift across areas should be examined further.

In terms of practical implications, we empirically uncover the crowd co-creation process in crowdsourcing. Firms which frequently crowdsource works to community members or seek crowd co-creation from a community should attract more experienced members in both diverse knowledge and specialized knowledge by increasing the variety and specificity of crowdsourcing tasks. Community members with only diverse experiences may not be very useful. Firms also need to attract those who have T-shaped knowledge in other areas by designing tasks with high relatedness and interactions along with a more flexible crowd co-creation process. Second, by attracting more experienced members and right type of members, firms could also spend less effort in assimilating the crowdsourced works because of better crowd performance. Third, it may be important for the designer of crowdsourcing tasks to understand the members’ balance between task exploration and motivation in crowdsourcing to guarantee the innovativeness and engagement of members.
References


Crowd Experience in Crowdsourced New Product Development


