Innovativeness of Ideas from Crowdsourcing Discussions

Albert Armisen Morell

DOCTORAL THESIS

Innovativeness of Ideas from Crowdsourcing Discussions

Presented by Albert Armisen

ESADE Business School

Department of Marketing Management, Operations Management and Innovation
and Information Systems Management

Supervised by Ann Majchrzak & Esteve Almirall

To my lovely wife and family

Acknowledgments

My journey started as a child when I was always wondering about the why of things and people's behaviour. The journey continued offering valuable experiences that I cluster into different stages. These stages allow me to recognize when I significantly evolve from a previous stage. However, I did not pursue this path alone because it was anonymous people at a time who guided me in this never-ending process. They are the igniters and force of change who sometimes illuminate alternative paths using their inner light and love.

Sheila Oller is the keystone who is always with me either on the brightest and darkest days. She balances and supports me on any type of day. I am enjoying the days and days to come with her by my side.

My parents, who both helped to discover and guided me during the early stages of my life, and are always offering emotional and rational advices. My grandmothers, who taught me about how to value what others do for you, and how others value you. My grandfathers, who understood my behaviour, style and space, and accepted me for who I was. My friends, who I admire by facing and overcoming the difficulties that life sets you. I would also want to highlight their importance to get me out of my comfort zone, despite it difficulty to do it.

Esteve Almirall who introduced me to innovation and gave me guidance about research and its implications. Ann Majchrzak mentored, guided me and helped to understand myself in perspectives I never considered.

Thank you all.

Abstract

Open innovation seeks to make organizations more competitive by managing the inflows and outflows of their organization. These organizations use crowdsourced discussions as a mechanism to encourage innovative ideas. However, the results of most of these crowdsourced discussions are not as innovative as expected. Some scholars seek to address this problem by mainly focusing on identifying appropriate incentives to encourage ideas, but they have not paid attention to how the structural characteristics of the crowdsourced discussion generate innovative ideas. This thesis draws from theories on management and creativity to develop a new framework about the effect of knowledge diversity on innovation in crowdsourced discussions. The proposed framework is validated using data obtained from multiple crowdsourced discussions. Moreover, a new methodology is presented for assessing the innovativeness of the ideas generated during crowdsourced discussions that is suited to the fuzzy nature of novel ideas. In sum, the three research questions addressed in this thesis are: 1) how does the diversity of the crowdsourced discussion affect innovativeness? 2) how do the characteristics of the crowdsourced discussion affect innovativeness? 3) how can the innovativeness of ideas be rated in crowdsourced discussion to take into consideration the different degrees of expertise from the decision-makers? The results indicate that less prior individual diversity, more current individual diversity, collaboration among participants, and first time posters foster innovation in crowdsourced discussion.

Contents

List of Figures

List of Tables

Innovation has nothing to do with how many R&D dollars you have. When Apple came up with the Mac, IBM was spending at least 100 times more on R&D. It's not about money. It's about the people you have, how you're led, and how much you get it.

Steve Jobs

1 Introduction

Firms might disappear if they are not constantly innovating (Tushman 1997). Some firms may seek help from the crowd to maintain a competitive advantage through consistent innovation (Easley and Kleinberg 2010; Surowiecki 2005). This thesis research examines how firms effectively produce innovative ideas and the role of diversity of the individuals participating in the crowdsourced discussions, which, is an essential element in innovation (Boudreau 2012; Brabham 2008).

This thesis explores the effects of the structural characteristics of the crowdsourced discussion on innovation by drawing from existing crowdsourcing theory, and presents theoretical and managerial implications. This thesis also proposes a new method to assess innovation when judges or reviewers have different levels of expertise in the ideas they are assessing by proposing a fuzzy method that captures this subjectivity more accurately using confidence degrees through fuzzy kappa measure.

1.1 Objectives and Research Problem

Crowdsourced discussions are critical mechanisms for opening organizations by incorporating outside ideas. Organizations should be able to manage external stakeholders to maximize the innovation potential. Despite the extensive literature on the factors that influence innovation in crowdsourced discussions (Bullinger et al. 2010; Füller et al. 2014; Hutter et al. 2011; Terwiesch and Xu 2008), these crowdsourced discussions have not led to ideas that are as innovative as desired (Linden 2014). Researchers have focused primarily on the study of the relevance and distribution of incentives and individual characteristics (Boudreau et al. 2011; Bullinger et al. 2010; Füller et al. 2014; Hutter et al. 2011); however, little attention has been given to study the structure of the of crowdsourced discussions. Thus, the main objective of this thesis is understanding the structural characteristics of crowdsourced discussions. Chapter 2 develops a theoretical framework of this thesis by analyzing the literature of the last ten years in crowdcourcing and later on narrowing to crowdsourced discussions. Chapter 3 theorizes about the role of diversity in crowdsourced discussions. Chapter 4 theorizes about different forms of participant contributions. Chapter 5 develops a measure to assess innovation that considers the varying degrees of confidence for each idea by the different decision-makers.

1.2 Contribution of Research

This thesis aims to contribute to the literature on crowdsourcing, innovation and diversity in order to develop a theory on the effects of structural characteristics of a crowdsourced discussion on stimulating innovation. The thesis identifies these characteristics in the discussion by distinguishing between the depth of the diversity of knowledge in the current discussion thread and previous discussion threads. Moreover, this thesis concludes that first-time posters and those participants seeking collaboration instead of argumentation have a greater innovation capacity. Finally, a new methodology for evaluating innovation that considers the different levels of expertise according to the decision makers was developed and validated.

The first step to this process was mapping out the literature in crowdsourcing, which was reviewed in Chapter 2. The theoretical framework lead to three main contributions: 1) How Diversity Contributes to Innovation in Crowdsourcing: an Evolutionary Diversity Model 2) Tapping the Innovative Business Potential of Innovation Contests, and 3) Consensus in Innovation

Contests Categorization by means of Fuzzy Partitions. A description of the contribution of each chapter is summarized below.

1.2.1 How Diversity Contributes to Innovation in Crowdsourcing: an Evolutionary Diversity Model

One of the most prominent assumption about the value of diversity in collaborative-based crowdsourced discussion is challenged in Chapter 3. This assumption is that innovation is more likely when the crowd is diverse because, as the online interactions occur, diverse knowledge perspectives are shared, new ideas are sparked, knowledge is recombined in new ways, and innovative solutions emerge. However, literature on creativity and innovation in teams suggest that diversity may not always have a positive impact on innovation. This chapter draws from the literature to hypothesize that the value of diversity depends on when it occurs during the crowdsourcing challenge. Theoretical implications for crowdsourcing, large-scale problem solving and organization design are that organizations should (1) seek individuals who have the creative capacity to integrate prior knowledge and (2) build upon the creative capacity of these individuals by helping these individuals join less diverse discussion threads, and later help them join highly diverse discussion threads. These findings suggest a notion of "evolutionary diversity" in crowdsourcing in which the discussion threads that people participate in affect them in ways not currently suggested by existing management theories.

1.2.2 Tapping the Innovative Business Potential of Innovation Contests

Few research attention has been paid to the structural characteristics of innovative crowdsourced discussion. Chapter 4 seeks to overcome this issue by developing three characteristics of innovative crowdsourced discussions. The first characteristic is where, in a discussion thread, the person should contribute given the discussion thread's amount of variety (i.e. variation of participants' familiarity with the topic or organizational background). Another characteristic is the amount of collaborative vs argumentative posts that have been made in the discussion prior to the person's innovative post. Finally, the last characteristic examined is whether or not the discussion includes posts from the same person prior to them submitting an innovative post. Chapter 4 uses data from an innovation contest. The analysis conducted found that there were three ideal profiles for generating innovative ideas: 1) persons benefit from posting in those crowdsourced discussions with substantial diversity, 2) persons offering the innovative idea should post in discussion

threads where participants focus their contributions on adding their perspectives not on arguing with others' perspectives, and 3) persons that contribute one-time to the to the crowdsourced discussion are more innovative than repetitive participants.

1.2.3 Consensus in Innovation Contests Categorization by means of Fuzzy Partitions

Finally, the task to assess innovativeness is overcomed by proposing a new measure of consensus among decision-makers to account for the fuzzy nature of innovation. Consensus decision-making rests on properly measured agreement. Chapter 5 proposes a fuzzy measure of agreement through fuzzy kappa based on fuzzy partitions. These fuzzy partitions enable decision-makers to assess their decisions with a degree of confidence. A fuzzy partition is built for each decision-maker by considering his/her self-assessed confidence degree when categorizing a set of alternatives or solutions. This enables decision-makers to capture the fuzzy nature of the decision more easily. In addition, this chapter uses data from five innovation contests to explore the viability of this new methodology based on using confidence degrees in real-life applications compared to traditional consensus decision-making. The results suggest that the use of confidence degrees improves the level of agreement in the consensus decision-making process through fuzzy kappa coefficients, and it also improves the level of agreement in the consensus decision-making process. Furthermore, the proposed method allows capturing the fuzzy nature of innovativeness by requesting fewer additional information to the decision-maker compared to previous fuzzy methods.

1.3 Structure of the Thesis

As detailed above, this thesis is a monograph on the basis of three articles and is organized as follows. Chapter 2 corresponds to a state of the art and theoretical framework of crowdsourced discussion that leads to three research questions. The objective in Chapter 3 is to study the evolution of diversity in crowdsourced discussions. The objective in Chapter 4 is to analyze the characteristics of crowdsourced discussions, providing managerial implications. The objective in Chapter 5 is to propose and validate a new methodology to measure innovation when there is low agreement among decision-makers. Finally, Chapter 6 offers an overview on the thesis with theoretical and managerial implications, limitations and future research.

Furthermore, a summary of the three chapters that addresses each of the proposed research questions (i.e. chapters 3, 4 and 5) is depicted on Table 1.3.1 decomposed into: conference or

journal, name, impact factor if any, research question, research setting, unit of analysis, research design, key findings, contributions, and implications. Each of this chapter addresses each of the research question raised in Chapter 2, and they also include a specific literature review about this research question.

	Chapter 3	**Chapter 4**	**Chapter 5**
Conference or Journal	Conference	Journal	Journal
Name	Organizing Crowds and Innovation	Business Horizons	Applied Soft Computing
IP	NA	1.008	2.857
Chapter Title	How Diversity Contributes to Innovation in Crowdsourcing: an Evolutionary Diversity Model	Tapping the Innovative Business Potential of Innovation Contests	Consensus in Innovation Contests Categorization by means of Fuzzy Partitions
Research Question	How does the diversity of the crowdsourced discussion affect innovativeness?	How do the structural characteristics of the crowdsourced discussion affect innovativeness?	How can the innovativeness of ideas be rated in crowdsourced discussions to take into consideration the differing degrees of expertise from the decision-makers?
Research Setting	Several crowdsourced discussions	One crowdsourced discussions about entrepreneurship	Several crowdsourced discussions
Unit Analysis	Post	Post	Batch or a set of posts
Research Design	Observational study	Observational study	Field experiment
Key Findings	Diversity has a time dimension for innovativeness	Three discussion characteristics foster innovativeness: Diverse people, Collaborative-based posts, and First-time posters	Time efficient methodologies to asses innovation taking the fuzzy nature into account are feasible.
Contributions	Provide a time based model to bridge the seemingly contradictory findings.	Participants should seek to collaborate rather than argue. Crowdsourced discussions benefit from a variety of participants to foster innovation. First-time posters are a valuable source of innovation.	Develop a soft group decision-making methodology based on confidence degrees.
Implications	Researchers should pay attention to the evolutionary diversity to asses the structure of the discussion.	Managers should attract as first-time posters from varied sources, while setting the crowdsourcing event to foster collaboration	Confidence degrees are an alternative for measuring fuzziness in group decision-making. Group decision-making techniques can be applied to measure innovation.

Table 1.3.1: Summary and findings of thesis's chapters

You don't have to be a genius or a visionary or even a college graduate to be successful. You just need a framework and a dream.

Michael Dell

2 Theoretical Framework

This chapter provides an overview of the theoretical background used to establish the foundations for Chapters 3, 4, and 5. This overview includes the relevance and difficulties of assessing innovation 2.1.1 and of understanding the structure of collaborative conversations in crowdsourced discussions 2.1.2.

2.1 State of the Art

Seven key areas were identified after selecting all the articles that mentioned crowdsoucing or crowdsourced that were published in peer-reviewed journals during the last ten years. There were a total of 2,067 articles that were used to determine these areas by using a topic analysis (i.e. LDA) following a semantic analysis method that allows identifying key topics (Evans 2014). This leads to seven different areas or topics, as described in the table: social media, collaborative-based crowdsourcing, practice, spatial crowdsourcing, methods, mechanical turk and competitive-based crowdsourcing. The frequency column in this table indicates the importance of this area in the selected articles. This thesis puts emphasis on collaborative based crowdsourcing.

Topic	Frequency	Keywords
Social media	7%	health, social, twitter, care, abstract, media, messages, tweets, risk, research, news, online, participants, findings, patients, positive, medical, attitudes, negative, effects
Collaborative-based crowdsourcing	22%	crowdsourcing, innovation, ideas, research, open, design, paper, article, participation, idea, translation, development, business, projects, crowd, community, model, findings, knowledge, creation
Practice	18%	social, data, crowdsourcing, research, technologies, public, science, media, citizen, web, management, government, citizens, challenges, resources, services, technology, issues, planning, open
Spatial crowdsourcing	10%	data, land, spatial, geographic, map, urban, mapping, maps, vgi, cover, osm, monitoring, road, location, geospatial, volunteered, areas, geo, locations, information
Method	10%	search, semantic, visual, annotation, data, clinical, image, images, web, video, game, annotations, approach, text, wikipedia, visualization, biological, techniques, information, drug
Mechanical Turk	14%	crowdsourcing, task, mechanical, turk, quality, workers, data, video, amazon, participants, tasks, results, performance, expert, responses, cost, mturk, crowd, agreement, assessment
Competitive-based crowdsourcing	20%	crowdsourcing, tasks, workers, proposed, problem, algorithm, mobile, propose, task, system, based, quality, learning, paper, model, real, human, show, data, algorithms

Table 2.1.1: Semantic literature review of 2.067 peer reviewed articles from 2006 to 2016

Crowdsourced discussions use crowds (Howe 2006) for different purposes such as: solving complex mathematical problems (e.g. TopCoder (Lakhani et al. 2010), financing new projects (e.g. KickStarter (Kuppuswamy and Bayus 2015)), working on long-term collaborative projects (e.g. Wikipedia (Leuf and Cunningham 2001)) or even short-term collaborative projects (e.g. BrightIdea (Tierney and Drury 2013)). Some short-term collaborative projects aim to generate ideas and select the most innovative ideas from a variety of sources (e.g. users, suppliers, employees, manufacturers, regulators) to identify new products, processes, ideas, or strategic individuals (Chesbrough 2006; Von Hippel 1989). These types of crowdsourced discussions are known as collaboration-based crowdsourcing (Afuah and Tucci 2012), henceforth referred to as crowdsourcing. These crowdsourcing events are used by governments (Administration 2 03) and non-governmental organizations (Malhotra et al. 2016), as well as by the private sector (King and Lakhani 2011; Malhotra and Majchrzak 2014).

As mentioned above, crowdsourced discussions are used to create new ideas through crowdsourcing events where participants voluntarily compete and collaborate to solve an organizational problem during a short period of time (Afuah and Tucci 2012; Bullinger et al. 2010; Hutter et al. 2011; Terwiesch and Ulrich 2009). This organizational problem is usually introduced in the form of *wicked* question with multiple alternative solutions (Rittel and Webber 1973) such as suggestions for new business models, new sources of revenue, or new strategic priorities (Majchrzak and Malhotra 2013; Malhotra and Majchrzak 2014). These types of problems are solved by the inner creative potential of the participants and also by the co-creation that is achieved in different discussion threads (Malhotra and Majchrzak 2014). Each discussion thread is composed of one or more posts generated by individuals who may contribute one or more posts in a crowdsourced discussion. The composition of participants in a discussion thread provides the diversity of the discussion thread. This diversity may be operationalized as the raw number of participants or the proportion of a feature of the participants such as level of expertise. Thus, this diversity fosters innovative posts in co-creation activities (Boudreau 2012; Boudreau and Lakhani 2015; Brabham 2008; Frey et al. 2011a; Howe 2006; Terwiesch and Xu 2008).

Despite the wide literature on factors influencing innovation in crowdsourced discussions (Bullinger et al. 2010; Füller et al. 2014; Hutter et al. 2011; Terwiesch and Xu 2008), these crowdsourced discussions have usually not led to ideas that are as innovative as expected and desired (Linden 2014). Researchers have mainly studied the relevance and distribution of incentives on crowdsourcing outcomes (Boudreau and Lakhani 2009; Bullinger et al. 2010; Füller et al. 2014; Hutter et al. 2011), but little attention has been given to the structure of the crowdsourced discussion as a factor which impacts outcomes. Moreover, an important point for the success of these crowdsourced discussions is knowing how to properly determine whether a post in a discussion is innovative or not. This categorization task poses several difficulties due to the fuzziness or subjectivity of innovation (Piller and Walcher 2006), the high number of posts to assess (Santos and Spann 2011), and the scarce time of those stakeholders who have the knowledge to evaluate the innovativeness of the posts.

2.1.1 Innovation

Businesses and organizations need to innovate through new products and services due to increased competition, globalization, and increased production costs. Innovation consists of creating new or improved services, products or production methods to redefine organizations and products

(McGrath et al. 1996; Schumpeter 2013). This thesis focuses on idea creation and draws from Amabile (1988)'s definition of innovation, whereby innovation is a recommendation from the crowd that is novel to, and implementable by, the organization (Amabile 1988). This definition of innovation is aligned with the paradigm of open innovation literature, where an organization judges the innovativeness of an idea relative to the existing business (Vanhaverbeke and Cloodt 2015).

Innovation in a crowdsourced discussion can occur after a recombination process in which the different knowledge areas of participants is integrated into a novel product, service or method (Dahlander and Gann 2010; Tsoukas 2009). The ideas produced during these crowdsourced discussions generate novel and average ideas in many fields, thus making novelty identification a challenge. The identification of these ideas is generally done by various stakeholders or decision-makers of sponsoring organizations. However, this process is more difficult due to the inability of decision-makers to properly identify novelty in the fields in which they are experts. This inability occurs because decision-makers assess harshly those ideas relevant to their own fields of expertise, while more leniently assessing ideas outside of their expertise (Boudreau and Lakhani 2015). For instance, while some decision-makers can assess an idea as innovative they might not be entirely confident because this particular idea involves knowledge that it is outside their knowledge expertise, other decision-makers might consider the same idea novel yet they might not be certain about it. Thus, each decision by a decision-maker or expert is done with a specific level of confidence. The expert opinion should be more highly regarded, yet traditional methods do not weight evaluations in this manner. Some authors suggested that new methodologies could be used in information systems based on decision theory, multicriteria decision making methodologies, fuzzy sets and modeling tool (Zaraté et al. 2008).

Wicked Problems: A Fuzzy Task

Problems that are assessed in crowdsourced discussions have not a unique answer and usually involve expertise from multiple areas because these problems are wicked problems. Innovation in wicked problems is measured at the organizational (Godart et al. 2014), individual (Bharadwaj and Menon 2000) and even discussion post level (Füller et al. 2012b). Innovation can be evaluated differently, through objective measurements such as number of patents (Wang et al. 2013), or more subjective measurements such as surveys (Füller et al. 2012b). In the case of crowdsourced discussions, a high agreement among decision-makers is generally used to deal with the subjectivity of the measures of innovation. However, as noted previously, different decision-makers have varying levels of confidence in the assessment of ideas in the crowdsourced discussion,

which is typically not considered when assessing the agreement among decision-makers. For example, consider a situation in which two decision-makers both evaluate an idea to be novel; one of these decision-makers is an expert in the field of the proposed solution he is evaluating and is certain about his evaluation, while the other is not an expert and is uncertain about his evaluation. The expert opinion should be more highly regarded, yet traditional methods do not weight evaluations in this manner. Therefore, assessing the innovativeness of wicked problems requires a fuzzy nature that takes into account the different levels of confidence from the evaluations of decision-makers.

This wicked task is done by using decision-makers to assess the innovativeness of the proposed ideas, their agreement can be measured in different ways, which is studied in group decision-making (GDM). GDM seeks to produce a final decision based on the selection of several possibilities from different stakeholders or decision-makers while measuring their agreement (Edwards 1977). The measurement of agreement is an important step in GDM because this measure is the output variable of the process (Herrera-Viedma et al. 2007, 2002). In particular, soft-group decision-making measures have the potential to capture the fuzzy or subjective nature of decision-making capturing the levels of confidence. These measures usually assess the agreement between decision-makers using the fuzziness between the preferences of two stakeholders or decision-makers (Herrera-Viedma et al. 2007, 2002). For example, a set of decision-makers seeks to classify a painting based on the century in which it was created by selecting one of three possible options (e.g. seventeenth, eighteenth, or nineteenth); so they are asked to rate how confident they are that the paining belongs to the specific century. The answer to these questions is used to assess the task in a fuzzy manner by considering the selecting and how confident they were at the time of the decision.

2.1.2 Crowdsourced Discussions

Crowdsourcing is "the act of out sourcing a task to a crowd, rather than to a designated agent (an organization, informal or formal team, or individual), such as a contractor, in the form of an open call"(Afuah and Tucci 2012, p.355). Crowds voluntarily commit to solve the task defined in the open call, although the organization can encourage their participation through various incentive mechanisms (Estellés-Arolas and González-Ladrón-de Guevara 2012; Howe 2006; Terwiesch and Ulrich 2009). There are two forms of crowdsourcing: tournament-based crowdsourcing, and collaborative-based crowdsourcing. Tournament-based crowdsourcing is based on each agent

competing independently against the crowd to develop the best solution possible; in contrast, collaborative-based crowdsourcing is based on allowing and fostering cooperation within the crowdsourcing event to reach the best possible solution (Afuah and Tucci 2012). This thesis focuses on crowdsourced discussions which are a form of collaborative-based crowdsourcing that last from three days to three months and may have intrinsic or extrinsic incentives (Bullinger et al. 2010). In particular, this paper will be examining collaborative-based crowdsourced discussions in which the sponsoring organization of the crowdsourced discussion makes an open call using a variety of sources such as emailing, newsletter or newspapers to attract a large crowd (Afuah and Tucci 2012; Boudreau and Lakhani 2009).

The organization, which does the open call, seeks to attract a large crowd hoping to tackle wicked problems by encouraging the creation of new ideas inspired by collaboration among the participants (Dahlander and Gann 2010). These crowds are characterized by participants who are: 1) seeking to generate innovative ideas, 2) volunteering their time to a problem of an organization and 3) seeking to receive rewards according to the design of the crowdsourced discussion (Afuah and Tucci 2012; Bullinger et al. 2010; Franke and Shah 2003; Leimeister et al. 2009; Terwiesch and Ulrich 2009). Furthermore, the main objective of these crowdsourced discussions is to foster collaboration, allowing participants to co-create innovative ideas (Majchrzak and Malhotra 2013; West and Bogers 2014). These collaborations among the participants are a key factor to generating the most innovative idea (Blohm et al. 2011; Bullinger et al. 2010; Franke and Shah 2003; Hutter et al. 2011). Collaborations take place when participants provide feedback by commenting or voting on others' ideas (von Hippel and von Krogh 2003) in a discussion thread.

The Structure of Crowdsourced Discussion and the Temporality of Diversity

The participants in a discussion thread interact with each other generating temporal groups (Ebner et al. 2009; Franke and Shah 2003; Jeppesen 2005; Piller and Walcher 2006; Sawhney et al. 2005; West and Lakhani 2008). Discussion threads are ad-hoc; temporary groups that arise during a crowdsourced discussion (van Osch and Avital 2010). These groups, whose members do not necessarily know each other before-hand (Faraj et al. 2011), facilitate the interaction and co-creation among the participants (Jeppesen 2005; Piller and Walcher 2006; West and Lakhani 2008). Therefore, these discussion threads are fundamental elements of crowdsourced discussions.

Crowdsourcing participants generally join several discussion threads during a collaborative-based crowdsourced discussion, where each discussion thread is influenced by chunks of knowledge from the individual participants. In management literature, the composition of the group

(i.e. participants in the discussion threads) allows to assess the team diversity, which has a positive relationship with innovation (Harrison and Klein 2007; Homan et al. 2008; Joshi and Roh 2009; West 2012). The positive effect of diversity in innovation has been found in stable and ad-hoc groups (Harrison and Klein 2007). However, crowdsourced discussion threads are dynamic groups that are constantly evolving over time. For example, two individuals in a discussion thread who have different backgrounds may not have the time to overcome the initial challenges of diversity because they might not be able to establish a shared language to successfully communicate with each other (e.g. an engineer and social scientist have different backgrounds, and therefore, different languages that if a shared language is achieved might lead to innovativeness). These types of concerns are not present in traditional teams (Harrison and Klein 2007) because the teams have a longer time to interact. Therefore, the theorized mechanisms on how diversity influences innovation might not be accurate unless time dimension is considered.

2.2 Research Gap

Scholars have paid significant research attention to online long-term collaboration among organizations (Laursen and Salter 2006) and communities such as Wikipedia (Hahn et al. 2008; Kane et al. 2014). However, there is a lack of research on the structural characteristics of online short-term collaborating crowds, such as crowdsourced discussions. Instead, most research focuses on encouraging innovation in these crowds through different methods such as using appropriate incentives (Boudreau and Lakhani 2009; Bullinger et al. 2010), ensuring the degree of cooperativeness vs. competitiveness that a participant (Hutter et al. 2011), or the design of the crowdsourced discussion (Alexy et al. 2011; Boudreau and Lakhani 2009; West 2009) should have to generate innovative ideas. One way to encourage collaboration is by incentivizing collective results (e.g. offering a prize for best collaborative spirit) instead of only incentivizing individual results (e.g. prize for best idea). Furthermore, when equal incentives for cooperation and competition are provided, people engage in behaviors that are both cooperative and competitive (Hutter et al. 2011). Another method to promote collaboration is by encouraging the use of non-technical language, proper terms and conditions that are aligned with the crowd, and defining objectives. These designs affect in turn the structural characteristics of the crowdsourced discussions (e.g. designing the incentives to foster cooperation might lead to more one on one discussions rather than engaging into argumentative discussions). However, this thesis focuses

on what is the ideal structure characteristics of the crowdsourced discussions on collaboration rather than how to design for them.

The first important characteristic of crowdsourced discussions is diversity from the individual familiarity with the topic (i.e. diversity of the crowdsourced discussion). This diversity is based on the different crowd composition of that an individual interacts with specific. Diversity is usually associated with positive effects on innovation (Boudreau 2012; Brabham 2008; Frey et al. 2011a; Howe 2006; Surowiecki 2005; Terwiesch and Xu 2008). Moreover, diversity in a crowdsourced discussion is associated with greater access to information (Page 2008) and access to different perspectives from the different individuals (Harrison et al. 1998; Taylor and Greve 2006), which in turn increases the supply of individual information. Therefore, diversity increases opportunities for co-creation between individuals because a participant does not have all the necessary knowledge to solve a problem (Bayus 2013; Malhotra and Majchrzak 2014). However, diversity presents challenges when integrating the diverse knowledge from several participants (Tiwana 2008), which cannot be done by everyone (Majchrzak et al. 2013). Participants who want to take advantage of the diversity to which they are exposed require some sort of semantic agreement (Carlile and Rebentisch 2003) and relational capital (pauluso3group). These two lines of thought raise the issue about whether diversity is good or bad for innovation in crowdsourced discussions, an issue which is addressed in Chapter 3.

Another important characteristic of crowdsourced discussions is related to the approach in the crowdsourced discussion. Participants can discuss in a breadth mode (i.e. commenting on the main idea) or depth mode (i.e. engaging in a one-on-one replies) during a discussion thread. Breadth mode implies that participants add their perspective to the main idea (Majchrzak et al. 2012a), in contrast to depth mode, which implies that participants argue about one specific issue (Tausczik et al. 2014). The literature suggests that the adding perspective (i.e. breadth mode) generates innovation by avoiding interpersonal conflict (Majchrzak et al. 2012a). This is In contrast to engaging in one-on-one discussion (i.e. depth mode), which allows ideas to be improved upon by exploring all the details. Thus, a dilemma arises: should the discussion thread have an additive or argumentative approach to encourage innovation outcomes? Therefore, an additional research gap is addressed in Chapter 4 to solve this point.

A different structure characteristic of crowdsourced discussions is based on the value of the first-time post of participants, either repetitive or one-time posters. Some scholars, as Feller et al. (2012), argue that repetitive, or core participants, are most likely to be brokers between different perspectives, and therefore, are more likely to reframe the problem to generate innovative ideas.

Other scholars, such as Faraj et al. (2011), propose that first-time posters spend considerable time observing, which gives them the opportunity to recombine different perspectives. So, the third research gap identified and discussed in Chapter 4 seeks to resolve the question of whether repetitive or first-time posters are the most innovative contributors in a crowdsourced discussion.

The final research gap that this thesis addresses is how to measure innovation. Measuring innovation should require capturing the degree of confidence of the expert who assesses an idea. Moreover, a measure that captures the fuzziness of decision-making might help to address a situation in which there is low agreement between decision-makers. This is usually achieved by capturing pairwise preferences of all options (Herrera-Viedma et al. 2007, 2014, 2002). This methodology requires that individuals invest significant time completing these comparisons. However, time is a significant resource and often a constraint for decision-makers, especially high profile individuals such as Chief Information Officers (CIO) (Toubia 2006). Moreover, crowdsourced discussions usually generate a large number of ideas (Diehl and Stroebe 1987; Neyer et al. 2009; Simonton 1999), which in turn increases the number of pairwise comparisons necessary. Finally, some methodologies exist that take into account the subjectivity of decision making, yet, these do not consider that some ideas are less likely to appear than others. Taking into account the proportion of idea is important because successfully detecting an innovative post from an idea that seldom appears is more important than detecting an innovative post from an idea that appears often. Therefore, Chapter 5 proposes a methodology that uses degrees of confidence while accounting for the different frequency distributions using a fuzzy index that takes into account the nature of innovation (Dou et al. 2007).

2.3 Research Questions

This thesis aims to understand the effect of the structural characteristics of crowdsourced discussions on innovation with the first two research questions. Moreover, a third research question evaluates the agreement between decision-makers, taking into account their level of confidence when assessing an idea. These research questions are as follows:

- How does the diversity of the crowdsourced discussion affect innovativeness?
- How do the structural characteristics of the crowdsourced discussion affect innovativeness?
- How can the innovativeness of ideas be rated in crowdsourced discussions to take into consideration the differing degrees of expertise from the decision-makers?

2.3.1 How Does the Diversity of the Crowdsourced Discussion Affect Innovativeness?

The existing literature on crowdsourcing usually argues for the positive effects of diversity on innovativeness in crowdsourced discussions (Boudreau 2012; Brabham 2008; Frey et al. 2011a; Howe 2006). However, literature about diversity in crowdsourcing is mainly based on the number of participants (Boudreau 2012; Terwiesch and Xu 2008). Therefore, the number of participants does not capture if the distribution but rather captures the depth or richness. For example, there may be a thousand participants who have the same knowledge expertise; therefore, there will be a low diversity but high richness. Moreover, the literature provides seemingly contradictory results because some argue that diversity increases or promotes innovativeness (Boudreau 2012), while others argue that diversity hurts innovation (Blohm et al. 2011; Girotra et al. 2010). Moreover, the literature on diversity is derived from stable teams that may not be directly applicable to online environments, since online participants only interact with others based on posts and comments (i.e. participants are unaware of the age or race of other participants that if noticeable might trigger faultlines in offline settings). Therefore, we seek to answer the question how does the diversity of the crowdsourced discussion affect innovativeness?

2.3.2 How do the characteristics of the crowdsourced discussion affect innovativeness?

This research question is addressed by raising three sub-questions related to the characteristics of the crowdsourced discussion affecting innovation. The first sub-question about the structural characteristic explores what is the right mix of expertise (Frey et al. 2011a; Howe 2006) based on participant familiarity with the topic (Leonard-Barton 1998; Schulze and Hoegl 2008). So the question is: does the variety of familiarity with the topic in a discussion thread effect if a person is posting an innovative idea? The second characteristic is based on how participants bring their perspective to the community, whether through an argumentative approach (i.e. engaging in a one-to-one dialogue) or additive approach (i.e. engaging in a conversation by adding different views to a discussion thread). Therefore, the second question is: how should participants engage in the conversation in a crowdsourced discussions to increase their innovative potential? Finally, the last characteristic is based on the level of commitment of the participants in the discussion, as some scholars argue that one-time posters are the most innovative (Kane et al. 2014), while others

argue the opposite (Bayus 2013). Therefore, the last question is: Does repeated contributions of an individual participant increase or decrease the innovativeness of an idea?

2.3.3 How can innovativeness of ideas be rated in crowdsourced discussions to take into consideration the different levels of expertise from the decision-makers?

The judging of ideas by decision-makers is a challenging and complex process to measure because it requires organizations to assess multiple decision-makers. However, these decision-makers do not usually have the same level of expertise, which is associated with assessing differing degrees of novelty of an idea (Boudreau and Lakhani 2015). Moreover, some scholars such as Amabile (1988) propose that novelty be decomposed into several components, which makes the process of evaluating innovativeness of ideas even more challenging because it requires a higher agreement among more than one component of innovation. These components include implementability, novelty, and feasibility, all of which are ill-defined by nature. Moreover, each decision-maker does not have the same level of confidence in assessing an idea (e.g. an engineer may be an expert on the feasibility of the solution, but might not have as much knowledge as a business manager regarding the novelty of the idea). This confidence intervals might be used to do a fuzzy evaluation of the ideas. Therefore, the question raised is how can the innovativeness of ideas be rated in crowdsourced discussions to take into consideration the different levels of expertise from the different decision-makers?

We need diversity of thought in the world to face the new challenges.

Tim Berners-Lee

3

How Diversity Contributes to Innovation in Crowdsourcing: an Evolutionary Diversity Model[1]

3.1 Introduction

Crowdsourcing is a practice used to seek and engage diverse individuals often external to the organization (Felin and Zenger 2011), allowing the organization to develop strategic directions such as identifying services and products desired by the public (Chesbrough and Appleyard 2007; Lakhani et al. 2010; Nickerson and Zenger 2004). In particular, this paper focuses on the use of crowdsourcing for open strategy formulation, which is a particular form of collaborative based crowdsourcing, henceforth refer as crowdsourcing (Malhotra et al. 2016). In crowdsourcing,

[1] A version of this chapter is under review in Journal of the Association for Information Systems, whose previous version was presented at the 2015 in Organizing Crowds and Innovation in University of Oxford, the 2013 Open and User Innovation Conference in Boston and the 2013 R&D Management conference in Salzburg by Armisen A. & Majchrzak A.

organizations broadcast a challenge question to different segments of the public (i.e. end-users, suppliers and employers) who voluntarily offer ideas (Jeppesen and Lakhani 2010). Crowdsourcing may be structured as competitive innovation tournaments or collaborative discussions (Afuah and Tucci 2012; Felin and Zenger 2014). We are focused on the latter type of crowdsourcing - online collaborative problem-solving discussions – and particularly the way in which they result in innovative solutions. In particular, crowdsourcing discussions that initially attract a crowd incentivize them to collaborate to solve a shared firm problem (Malhotra et al. 2016). This anonymous crowd, through repetitive interactions with the members, becomes an initial stage of community that seeks to generate innovative ideas in a short period of time. Note that some authors as Villarroel et al. (2013) identified discussion threads closer to the community than to crowds. Following Amabile (1988), we define innovativeness as a recommendation by a member of the crowd that is novel to, and implementable by, an organization. This definition aligns with other scholars such as Lukyanenko et al. (2014) who define innovativeness as information quality, which is fitness of data for use by information consumers for specific purposes. The definition also aligns with the definition of innovation in the open innovation literature, as an idea that the organization judges to be helpful in identifying ways to foster existing business (Von Hippel 2005) or create early stage ventures in a new business area (Vanhaverbeke and Cloodt 2015).

A key determinant of innovativeness from crowdsourced discussions found in the literature is the diversity that the crowd brings (Boudreau 2012; Frey et al. 2011a; Girotra et al. 2010; Terwiesch and Xu 2008). Diversity refers to differences in experiences that lead to cognitively unique mental models of the world and the factors that influence the world (Harrison et al. 1998). This diversity stimulates innovativeness because each individual brings different knowledge which may stimulate others. Furthermore, when problems are wicked, as they are with open strategy formulation, require knowledge that is more diverse to solve them (Carlile and Rebentisch 2003). This diversity from the crowd is being addressed fostered via an open call, which is a critical element in the crowdsourced discussions (Estellés-Arolas and González-Ladrón-de Guevara 2012).

Most literature argued that crowdsourcing generally for the positive effect of diversity on innovative outcomes (Boudreau 2012; Brabham 2008; Frey et al. 2011a; Howe 2006). However, the results are not conclusive because diversity is not always significant (Blohm et al. 2011; Girotra et al. 2010). In one study by Girotra et al. (2010), for example, it was found that diverse crowds engaging in discussions with each other do not produce higher quality ideas because the crowd tends to ignore the diversity. In another study, diversity did not produce more innovative ideas

for a complex problem (Kavadias and Sommer 2009). As yet another example, Blohm et al. (2011) found that the discussions in a crowd have no effect on the average quality scores of ideas offered, including innovativeness. However, crowdsourcing literature assumes that diversity is static and does not change over time (Boudreau 2012; Terwiesch and Xu 2008), yet the diversity in crowdsourced discussion changes through time because new discussion threads are constantly generated. Therefore, the gap we address is to theorize, and provide empirical evidence, that may explain this contradictory relationship between diversity and innovation by considering time. In particular, we ask the following research question:

How does diversity affect innovativeness of ideas suggested in discussions during crowdsourcing?

By studying the role of diversity in crowdsourcing, particularly in crowdsourcing discussions for wicked problems, the theory about the role of diversity in large-scale problem-solving and organizational design (Ancona et al. 2001; Puranam et al. 2013) is enriched with an evolutionary perspective on effects of diversity that takes into account time. Moreover, the claim from Afuah and Tucci (2012) that the crowd's diversity cannot help with ill-structured problems can be overcome.

3.2 Conceptual Development

In this section, we first develop the different approaches conceptual development of diversity. We then explain the reasons why diversity has been assumed to have a positive effect on innovation in crowdsourcing, and then the reasons why diversity may have negative effects. We then propose a time perspective on diversity that allows both positive and negative effects to occur. Finally, we use that perspective to develop three hypotheses about the role of diversity on innovativeness of ideas offered in discussions during crowdsourcing.

3.2.1 Diversity

Crowdsourced discussions are user-generated content that organizations obtain by engaging a voluntary crowd to solve a wicked problem in order to attract a diverse crowd. While there have been many definitions offered for diversity, the most prevalent definition of diversity used in crowdsourcing literature is variety (Brabham 2008; Surowiecki 2005) or a proxy to variety through richness or raw number of individuals (Boudreau 2012; Terwiesch and Xu 2008). Variety is defined as the degree to which attributes are evenly distributed among the team members. For example, assessing diversity as richness supposes that nine expert individuals are more diverse

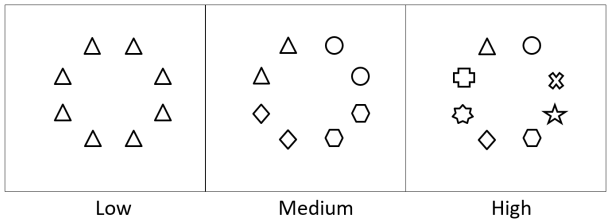

Figure 3.2.1: Graphic Example of Diversity as Variety where Each Symbol is an Individual and the Different Indicate Individuals with Different Attributes.

than three expert individuals; in contrast to considering high diverse as three sets of low, medium and high levels of expertise compared to a low diversity when having nine expert individuals. Figure 3.2.1 shows an adaptation of the Harrison and Klein (2007) depiction of low, medium and high amounts of variety.

When problems are wicked, as they are with open strategy formulation, the individual attributes needed to solve the wicked problem require knowledge of the topic but also having the capacity to think outside the box. This capacity to think outside the box is easier to those that are novices or not expert in the domain because they are unaware of path dependencies. These novices may see the problem from an unconventional perspective that might inspire experts, while novices might take inspiration from experts. Therefore, the expertise of the individual is a key attribute in crowdsourcing (Afuah and Tucci 2012) that should be assessed for diversity as others as done (Dahlin et al. 2005; Van Der Vegt et al. 2006). This diversity on the levels of expertise of the crowd capture the varying knowledge and heterogeneity of them related to the problem.

3.2.2 Positive Effects of More Diversity

Many crowdsourcing scholars have argued for a positive effect of diversity (Boudreau 2012; Brabham 2008; Frey et al. 2011a; Howe 2006; Surowiecki 2005; Terwiesch and Xu 2008). There are many reasons given for the positive effects of diversity on innovation. Diversity within groups allows access to a wider knowledge base and greater information richness (Page 2008). Individuals with different types of knowledge will have different knowledge to contribute to the team, providing different perspectives needed for innovation (Harrison et al. 1998; Jackson et al. 1995; Taylor and Greve 2006). Diversity also creates an opportunity for co-creation between the individuals because a single individual is unlikely to have all the required knowledge to offer innovative solutions to complex problems (Bayus 2013; Felin and Zenger 2011; Hargadon

2003; Majchrzak et al. 2004; Malhotra and Majchrzak 2014). Diversity fosters the discovery of complementary and unique information that can be recombined in new ways (Boudreau 2012; Felin and Zenger 2011; Tsoukas 2009).

In a discussion thread carried out during an crowdsourcing, the contributions made by the participants are typically organized into threads, with a discussion thread started by a question or comment or idea, and the contributions that follow are comments (Armisen and Majchrzak 2015). In such a discussion thread context, the positive effect of diversity may function in the following way. Although several incentives and designs are used to guide these crowdsourcing events (Hutter et al. 2011), an individual self-selects when they choose to contribute (Felin and Zenger 2011). Each knowledge exchange with others in the crowdsourced discussion tends to occur within discussion threads. Each thread has a set of posts that are read before an individual offers his/her own idea. The concept is that these posts in the discussion thread influence the person with the idea since they form an 'information supply' (Boudreau 2012; Terwiesch and Ulrich 2009; Terwiesch and Xu 2008). In a discussion thread, individuals with diversity may have been posting to the discussion thread, thereby expanding the information supply available to the individual. This expansion can challenge the individual's preconceived notions about the domain (Boland Jr and Tenkasi 1995). In an effort to reduce the cognitive dissonance associated with disconfirming information, the individual will conceptually expand or reframe his perspective, leading to an innovative idea (Tsoukas 2009).

3.2.3 Positive Effects of Less Diversity

The positive effect of diversity has detractors. Diversity creates the need to integrate different individuals' knowledge (Gray 2000; Tiwana 2008). However, not everyone has the ability to do this knowledge integration (Majchrzak et al. 2013). The integration of diverse perspectives requires a shared agreement across interdepend groups that may not be present in them (Carlile and Rebentisch 2003). This shared agreement increases the more diverse the groups because it increases the amount of dependencies that must be integrated to generate the final solution. Therefore, diverse groups require a higher cognitive capability to overcome this limitation, in contrast to less diverse or homogeneous groups that do not need to create this shared agreement. Furthermore, diversity may keep individuals from having sufficient relational capital (Paulus and Nijstad 2003) (i.e., trusting relationships with each other) to invest their time and energy to

establish a shared language for the solution (Carlile 2002). Diversity may lead to an inability to agree on the same goal, and thus an inability to co-create (Pieterse et al. 2011).

Thus, less diversity may foster innovation. Individuals in the crowdsourced discussion in discussion threads with less diversity may feel more motivated to offer innovative ideas since innovation is a shared goal. They may feel more confident in the ideas they have because they are cognitively similar to others that have been shared (Tsoukas 2009). Moreover, they may feel less concerned about personal risks of embarrassment because they trust others in the discussion thread (Edmondson 1999). They may be better able to build on others' ideas to conceptually expand the ideas suggested because of the language shared (Tsoukas 2009). Finally, they may be more willing to offer constructive advice on how to develop others' ideas further because of relational capital shared (Tiwana and Mclean 2005).

3.2.4 Effect of Diversity on Innovation May Be Based on Evolution over Time

Both perspectives on the effect of diversity may be accurate. Whether more or less diversity is useful may depend on when, during the crowdsourced discussion, the diversity is experienced.

The literature supporting the value of diversity for innovation typically refers to diversity that occurs as the innovative idea is being formulated and then shared. For example, innovation through dialogue is focused on individuals challenging each other to understand each other's perspectives; as individuals attempt to synthesize across the divergent perspectives, new ideas emerge (Gray 2000; Tsoukas 2009). Carlile (2004) has argued that innovation arises at the juxtaposition of multiple disciplines as individuals attempt to reconcile the differences such as the ones from novices and experts. As unique knowledge surfaces, it is during the attempt to integrate the knowledge that cognitive dissonance is experienced, causing a stress reaction; the individual will be motivated to take immediate steps to reduce the stress of dissonance, such as by integrating the knowledge in new ways. Without reduction of the stress, the individual will cognitively or physically withdraw from the discussion (Claxton 2000; Paulus and Nijstad 2003).

Therefore, we suggest that diversity is likely to have its greatest positive effect on innovativeness when the diversity occurs among the posts in the discussion thread immediately prior to the individual posting his or her idea. When the posts in the discussion thread are diverse, they create the cognitive dissonance, motivation, and catalyst to offer an innovative idea. When the posts in the discussion thread are less diverse, there is less cognitive and psychological stimulation to conceive of an innovative idea. Thus, we hypothesize:

H1: An individual is more likely to offer an innovative idea in discussion threads with a greater diversity of posts immediately prior to the individual's posting.

The literature on the value of less diversity on innovation may also be a function of time – but a different point in time. Much of that literature discusses the value of homogeneity in terms of building "creative capacity", the ability of an individual to recognize the value of new information, assimilate that information, and apply that information to create novel knowledge (Cohen and Levinthal 1990). Individuals learn how to address complex problems collaboratively when they are able to take advantage of the easiness to cooperate with those with similar knowledge (Locke and Horowitz 1990). Edmondson (1999) refers to psychological safety as a capacity of cooperation that allows an individual to take risks. Tsoukas (2009) refers to an inter-subjective meaning between individuals that is needed before they can co-create. Nahapiet and Ghoshal (1998) theorize about the importance of relational capital that is needed as a prerequisite for intellectual capital creation. Finally, Carlile (2004) refers to the semantic agreement among participants as a prerequisite for knowledge transformation. In sum, the argument that less diversity leads to more innovation can be conceptualized as a creative capacity-building construct. Such a capacity is determined during the discussion thread by the manner in which others react to the individual's posts. If a discussion thread of posts indicates similar language to that of the individual entering the thread, as well as a tolerance and encouragement for risk-taking, relational trust, and a general feeling of cohesiveness, then an individual's creative capacity may be developed by that thread. Such a capacity may need to occur prior to the dialogic process in which innovative ideas emerge. During crowdsourced discussions involving discussion threads, then, individuals may need to develop a creative capacity within the challenge first, before they can offer innovative ideas later. Since, in accordance with Hypothesis 1, diversity in a thread stimulates innovative idea generation, the lack of diversity within a thread is unlikely to generate new ideas during that thread; instead, the lack of diversity is likely to be helpful when it occurs prior to a stimulating dialogic process occurring in a follow on thread. That is, we suggest that the creative capacity has spillover effects in later high-diversity threads. That is, when an individual initially enters a crowdsourced discussion, she may need the low-diversity thread to build her confidence that others understand and respect her perspectives, and feel psychologically safe enough to share risky ideas. This capacity emboldens the individual as she moves to a new discussion thread. Even though this second thread has much more diversity, the spillover effect of her confidence and safety allows her to engage in challenging dialogue that stimulates new ideas. Thus, we hypothesize:

H2: The less diversity displayed in initial threads (called prior threads) that an individual enters, the greater the likelihood that an eventual post made by the individual will be innovative.

When less diversity occurs initially upon entering the challenge, and more diversity occurs later in the challenge, the effect on innovation should be stronger than when either more or less diversity occurs without the other and at different times. The failure to build a creative capacity early on in the challenge may constrain the individual from taking advantage of the diversity in a thread. An individual with little self-confidence is likely to not challenge others that, and thus fail to become sufficiently engaged in the dialogue to be challenged herself. Moreover, an individual feeling unsafe to share risky ideas is unlikely to offer innovative ideas, and gain the feedback from others to make the ideas even more novel and implementable. Lastly, an individual entering a crowdsourced discussion without experiencing the cohesiveness that prior threads could create may not feel wanted, respected, or encouraged to share, limiting one's contributions to small incremental suggestions or safe comments on others' ideas. Therefore, we offer our final hypothesis:

H3: The more an individual has been engaged in a prior thread with less diversity, the greater the probability to offer an innovative idea in discussion threads with a greater diversity of posts immediately prior to the individual's posting.

3.3 Methods

3.3.1 Data and sample

To test our hypotheses, we selected an organization sponsoring multiple crowdsourced discussions in which for each, multiple discussion threads were involved. A single organization was preferred to allow for control of extraneous factors such as type of organization, previous experience of the organization with crowdsourcing, public perception of the organization, and the nature of the public invited. In addition, having a single organization allowed for controlling for differences in how innovativeness of ideas were assessed since all ideas would be assessed by the same individuals. Finally, we wanted an organization in which multiple crowdsourced discussions were conducted to assess generalizability. The organization we selected was the United State Agency for International Development (USAID). USAID had conducted a crowdsourced discussion in six different areas in 2010 to help the organization define its strategic priorities for the following five years. The crowdsourced discussion focused on six different areas: (1) Fostering Science, Technology and

Innovation, (2) Building Stronger Partnerships, (3) Empowering Women and Girls, (4) Pursuing Grand Challenges, (5) Inspiring a New Generation and (6) Supporting a Sustainable Planet. The crowdsourced discussion was documented in detail in a USAID report (Ferguson 2010). The platform on which the crowdsourced discussion was conducted was organized around discussion threads, using the same platform on which successful IBM innovation jams have been conducted (Bjelland and Wood 2008). These six topics in the crowdsourced discussion together led to 988 threads, consisting of a total of 5.829 posts made by 1.662 individuals from 43 different countries. The analysis was restricted to threads with three or more posts in order to account for diversity within the thread; in addition, the first thread was not used so that prior threads could be examined. These restrictions reduced the number of threads analyzed to 423 threads, consisting of a total of 4.628 posts made by 1.411 individuals.

Each call for a topic was preceded by substantial marketing efforts of globally located local organizations funded by USAID. The first post in each crowdsourced discussion was made by a highly reputable person in the field (e.g., global director for public sector industry from Deloitte, ambassadors from different countries, professors from different universities such as Harvard University and Georgetown University, executives from NGO and private companies). Any participant started other discussion threads in the different crowdsourced discussions. The crowdsourced discussions were conducted only for 72 hours – all at the same time – to ensure intensity and focus of efforts during the crowdsourced discussion. For each crowdsourced discussion, participants completed a background profile when they registered using a real or fake name, making their backgrounds available for analysis. This background information was not available to other participants. There was neither voting on posts nor awards, so that participants had no disincentives to participate in discussion threads. Moreover, since participants were drawn globally and represented 43 different countries and hundreds of locations within each country, and many participants only used their first name or fake name, there was minimal possibility that participants knew each other.

3.3.2 Measures

The operationalization of control, independent and dependent constructs are discussed below.

INDEPENDENT VARIABLES

Diversity was measured by asking participants how familiar they were with the topic of the crowdsourced discussion (i.e. one of six topics) in which they were contributing (0= "Not

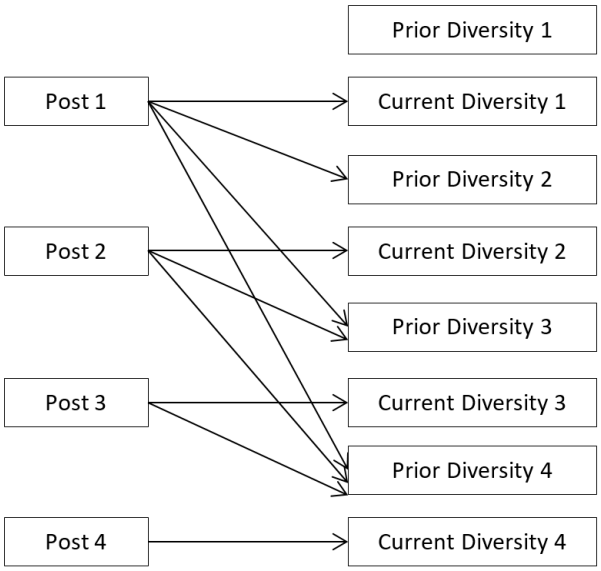

Figure 3.3.1: Use of Discussion Threads of and Individual for measuring his Current and Prior Diversities.

familiar", 1= "Just a little familiar", 2= "Somewhat familiar", 3= "Very familiar", 4= "I have worked in/presently work in this field"). Diversity was calculated for the threads in which each individual participated in. Thus, for an individual who participated in three discussion threads, diversity was calculated for each thread, resulting in three diversity measures. Then, the threads were distinguished into those which were referred to as "current" and "prior". A "current" thread was defined as a thread in which the individual had posted an idea that was evaluated for its innovativeness (see below how we evaluated innovativeness). "Prior" threads were all those threads in which the individual had posted prior to the current thread. If an individual had not previously posted, the individual was eliminated from the analysis so that the relative effects of diversity for both prior and current threads could be assessed. Any single individual could have multiple current and multiple priors depending on the number of posts they made. For example, an individual with three posts would have three current diversity and two prior diversity measures: the first post would have just current diversity, the second and third post would have a prior and current diversity measure. This process is explained in Figure 3.3.1 that also includes how the different discussion threads are reused. Furthermore, the multiple posting by individuals is taken into account in the analysis using a multilevel model.

Teachman (1980)'s formula for measuring variety was used to measure diversity for the current

thread and all priors put together. More precisely, to measure diversity in current threads, a small team version[2] of Teachman (1980)'s formula was used to assess the proportion of group members with each level of expertise from the current thread. To measure diversity in the prior threads, the same small team version of Teachman (1980)'s formula was used, with the proportion assessed based on all the individuals that the individual interacted with prior to the current thread. Therefore, prior and current diversity were a continuous variable that assess the homogeneity (i.e. negative value) and diversity (i.e. positive values). The current diversity is measured based on the immediate or focal discussion thread, while the prior diversity is based on the prior discussion threads the individual participated in.

DEPENDENT VARIABLE: WHETHER POST IS INNOVATIVE OR NOT

The 4.628 posts from these six topics in the crowdsourced discussion were reviewed by a committee of USAID representatives after the crowdsourced discussion were completed. For each of the six topics, the committee identified different strategic directions reported to be "inspired by" the posts that were considered as innovative to USAID (defined as novel and useful to USAID). The total of 40 strategic directions (five to eight strategic directions for each of six was described in an internal document (Ferguson 2010). These research considered innovative those idea that were novel according to the organization (Amabile 1988), which ensures a fitness for use of data by the organization (Lukyanenko et al. 2014).

Interviews with USAID staff indicated that the committee had not conducted a systematic analysis of the innovativeness of the 4.628 posts, but instead had simply read all the posts in each topic in the crowdsourced discussion, and attributed the strategic directions to the inspiration that the posts provided. Since the report did not specify which posts corresponded to which of the 40 innovative strategic directions, we needed to conduct a detailed semantic analysis to measure the innovativeness of each post, i.e., the semantic similarity between the post and one of the 40 strategic directions for that crowdsourced discussion.

In a process similar to semantic analysis (Goddard 2011), the strategic directions were parsed by identifying from one to three concepts that described each strategic direction. Table 3.3.1 shows in Column 1 the five innovative strategic directions identified for the Support a Sustainable Planet challenge. By examining the description of each strategic direction provided by USAID (Column 2 of Table 3.3.1), the key concepts associated with the description were identified (Column 3 of Table 3.3.1) we created a list of concepts. Coders not involved in the crowdsourced discussion

[2] $H_b = \frac{\ln N! - \sum(\ln n_i)}{N}$, where N is the total posts and n_i is the total post from knowledge $type_i$

or familiar with the research hypotheses were asked to identify (yes or no) if the concept existed in each post for that topic. Two raters completed the coding, achieving a percentage agreement of .95 ($\kappa = 0.93, z = 71$) using a squared weighted Cohen's Kappa (Gamer et al. 2012). When a post had one or more of the concepts in the description of the strategic direction, the post was referred to as innovative (1 vs 0 for not innovative). Of the 4628 posts, 447 were judged to be related to one of the 40 innovative strategic directions. Examples of innovative posts are shown in Table 3.3.1 (column 4) for the Sustainability Challenge. The remaining posts 4181 were, by default, rated as non-innovative.

CONTROL VARIABLES

The control variables and their definitions are shown in Table 3.3.2.

3.3.3 Analytical Strategy

Posts are nested within threads which are in turn nested within each crowdsourced discussion. These posts were generated by different individuals that could potentially contributed to different threads in different crowdsourced discussions. This creates a crossed multilevel data structure with three levels: a level 0 (i.e. post level), level 1 (i.e. thread) and two level 2 (i.e. crowdsourced discussion and individual posts) as shown in Figure 3.3.2. Thus, posts belong to discussion threads which are nested in crowdsourced discussion, as well as, individuals. For example, an individual generates four posts that belong to three different discussion threads, which in turn belong to two crowdsourced discussions. These types of crossed data are usually analyzed using a multilevel models that allows for capturing the similarity among posts contributed by the same individual, or within the same thread or crowdsourced discussion (Hox et al. 2010). The sample size at the post level was 4628 (level 0) in 423 discussion threads (level 1) and six crowdsourced discussions generated by 1411 individuals (level 2).

We used an R software library for conducting the crossed multilevel logit regression (Bates et al. 2014). The random intercept model in which the intercepts vary for each level 1 allowed for capturing the similarity within individuals, discussion threads and crowdsourced discussions (Hox et al. 2010). A multilevel model has two statistical components: fixed effects (i.e. regression parameters) and random effects (i.e. variance components). Random effects are assessed through F-test statistics, and fixed effects assessed through t-test. The fit of an HLM model is assessed in two different forms: (1) through the deviance (i.e. $-2 * \Delta \log \textit{likelihood}$) and its associated ϕ^2, and through (2) Akaike information criterion (AIC) that measures the quality of the model

1) USAID Innovative Strategic Direction	2) Description of Strategic Direction	3) Key Concepts in Innovative Strategic Direction	4) Example of a Post Coded to be Innovative Because of Its Similarity to Key Concept in Column 3
Develop tools that map supply and demand within communities.	Give preference to local products and leverage group demand for products not available locally.	Develop tools; map supply and demand within communities.	Solidarity economy is a possibility. In Northeastern Brazil, a Solidary Popular Bank was created, mapping supply and demand for products and services within a community/neighborhood. This means that preference is given to local products and services, keeping the local wealth local [...]
Provide international environmental education.	Intergenertaional focus on protecting the environment, integrate environmental action in an educational context.	Intergenerational reeducation Environmental Education.	[...] we try to aware people about what actually is ecology and how should we behave. We are working mostly with children and projecting ecological trips with multiple including of different topics. So, I can say that its working because we see that children then teach adults. It's amazing!
Promote eco-cities that show eco-firendly policies and economic growth can work in partnership.	Leverage sustainability plans developed at the scale of a city.	Promote eco-cities; eco-firendly policies; economic growth; partnership with eco-friendly policy and economic growth.	Cities, as it was mentioned before, aren't the problem but part of the solution. Concentrating the atmospheric gasses in one area makes it easier and even economically sound to intervene and reduce them in a bigger scale with better results and impact. [...]
Build capacity to enforce environmental responsibility in the private sector and the community.	Environmental exploitation often is not prevented due to a lack of laws or rules but a lack of government oversight and enforcement.	Enforce environmental responsibility in the private sector; enforce environmental responsibility in the community; lack of government oversight and enforcement on environmental issues.	[...] States must provide a space for communities to use their self-governing capabilities and other sources of resilience to adapt to their changing priorities. Resilience remains a key feature of adaptation. Governance is a tool for building resilience; failure to implement and enforce zoning and environmental regulations, for example, often results in development that increases vulnerability to climatic stresses. Many countries have national environmental plans, but do not enforce them. In addition, potential negative impacts exist at the nexus of land tenure, property rights and GCC. Judicial systems for land claims processing, prosecution of illegal logging, and other land use crimes are critical for militating against potential conflict.
Create small farm adaptation initiatives managed by farmer associations/communities for piloting adaptation strategies.	Engage small-hold farmers to understand the impact of climate change in their setting. Engage them in the investigation of adaptive strategies.	Engage small-hold farmers to understand the impact of climate change in their setting; investigate adaptive strategies for small farms; pilot adaptive strategies for small farms; manage initiatives by farmers' associations or communities.	[...] It's true that pastoralists draw on a wealth of ancestral knowledge on climate fluctuations in their region, and have often developed ways to cope. How then can we help these pastoralists anticipate future fluctuations? One way that USAID works to help developing countries anticipate these changes is through forecasting technology systems, such as the Famine Early System, and SERVIR. [...] How can we bring governments on board to enact policies and reforms that are more sustainable and account for climate change? [...]

Table 3.3.1: Top Five Innovative Strategic Directions For *Supporting A Sustainable Planet* Extracted From The USAID Committee Report.

Name	Description
(1) Expertise of the individual	Individuals who have a lack of domain expertise are not going to be able to know what the firm needs to, therefore produce less innovative posts.
(2) Average expertise of the focal thread	This is how close these individuals are to the firm's problem (Afuah and Tucci 2012). Moreover, this control variable allows capturing the ordinality nature of expertise. It is measured by the standard average expertise of all the individuals in the current thread.
(3) Average expertise of prior threads	An additional control variable to assess how close individuals are to the firm's problem (Afuah and Tucci 2012) is measured by the standard average expertise of all the individuals in individual interacted with in prior threads.
(4) Engagement of the individual	Idea generators may harm innovation since they often do not interact with other individuals (Füller et al. 2014). This variable is measured as the number of post an individual does.
(5) New knowledge produced	An individual might produce new novel knowledge compared to what he or she has produced before. This is measured by using a distance measured based on the topics of the post he produced compared to the ones in the focal post, following a semantic analysis procedure (Evans 2014).
(6) Cross-poster	An individual who bridges different communities is likely to be less innovative (Lim et al. 2010). This is measured as a dummy variable to determine those that contributed to more than one topic in the crowdsourced discussion.

Table 3.3.2: Control Variables and Their Associated Description

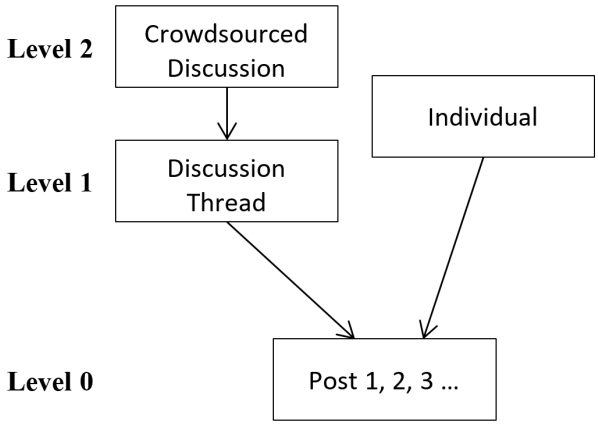

Figure 3.3.2: Description of the Different Levels for the Crossed Multilevel Model.

based on information theory (Bozdogan, 1987). Furthermore, we provided a conditional R^2 that accounts for the model predictiveness. To test our hypotheses, a control model was first assessed which consisted only of the different control variables: (1) expertise of the individual, (2) average expertise of the focal thread, (3) average expertise of prior threads, (4) engagement of the individual, (5) new knowledge produced, and (6) cross-posters. To test for Hypotheses 1 and 2, the less prior diversity and more current diversity for an individual's post was included. To test for Hypothesis 3 using HLM, the interaction effect of less prior diversity with more current diversity for an individual's post was included.

3.4 Results

The individuals participated by generating 10.1 posts producing knowledge that was less new than usual (0.42). The different levels of expertise from the individuals who participated in the crowdsourcing discussions were: 77 that were "not familiar", 171 that were "just a little familiar", 333 that were "somewhat familiar", 259 that were "very familiar", and 571 that answered "I have worked in/presently work in this field". The average level of expertise for those participating was "very familiar". The average expertise of individuals contributing to current and prior threads was "very familiar" (2.81 and 2.79 respectively). Table 3.4.1 describes the sample of individuals in the crowdsourced discussions according to the different types of expertise. There was no sign of multicollinearity considering that the highest correlation was $-.43$.

As explained above, the hypotheses were tested with multilevel logistic regression model with

	μ	σ	(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable:										
(0) Whether Post is Innovative or Not	0.07	0.26								
Control Variables:										
(1) Expertise of the individual	2.81	1.24	.04							
(2) Average expertise of the focal thread	2.80	0.54	.03	.42						
(3) Average expertise of prior threads	2.79	0.23	.01	.14	.38					
(4) Engagement of the individual	10.1	25.4	−.05	−.03	−.09	−.07				
(5) New knowledge produced	0.42	0.51	.04	−.03	−.02	−.00	−.43			
(6) Cross-posters	0.55	0.50	−.02	−.08	−.06	−.01	.41	−.41		
Independent Variables:										
(7) More current diversity	0.94	0.42	−.01	−.08	−.24	−.14	−.05	.12	−.01	
(8) Less prior diversity	1.33	0.24	.04	.04	.12	.43	−.06	.01	−.01	−.42

Table 3.4.1: Description and Correlation of the Variables.

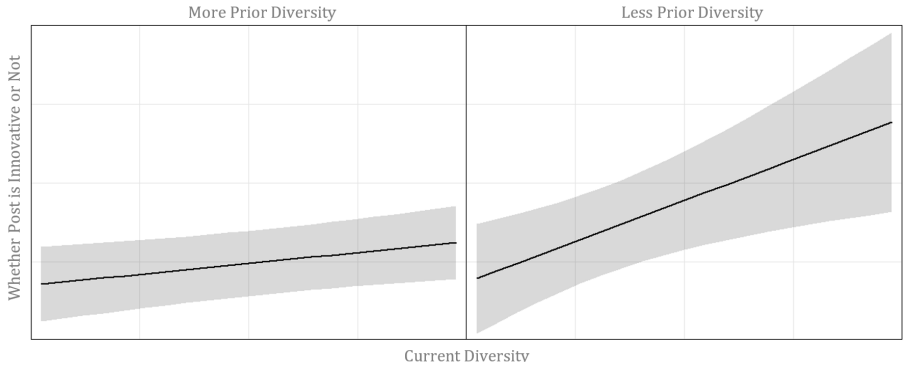

Figure 3.4.1: Interaction Effect on Diversity

post at level 0, and discussion thread at level 1, and crowdsourced discussions and individual at level 2. The model was tested against a model with control variables only. The results are shown in Table 3.4.2. All three hypotheses are supported such that the main effects and the interaction effect are significant, and contribute a significant increase in variance in the dependent variable over the controls ($\phi^2 = 11.8^{**}$). Finally the effect of individuals contributing in multiple crowdsourced discussions was examined by including in the analysis whether the individual had posted in other crowdsourcing discussions (i.e. cross-poster) which had no relationship with whatever the post was innovative or not.

Figure 3.4.1 displays the interaction effects between current and prior diversity. The straight line is the regressed coefficient, while the grey area is the 95% confidence interval. Apparent from the figure is that the interaction effects function as hypothesized; that is, the positive effect on innovativeness of diversity in an individual's current thread is increased when that individual's prior threads have less diversity. Thus, the more homogeneous the prior threads are and the more diverse the current thread is, the more innovative the post.

3.5 Discussion

The research question that was raised at the beginning was: *How does diversity affect innovativeness of ideas suggested in discussions during crowdsourcing?* The findings indicate that diversity in threads has both a positive and negative effect on innovativeness of posts depending on temporal order of the discussion threads. A greater number of innovative ideas come from individuals who enter crowdsourced discussions initially into less diverse discussion threads, followed by more

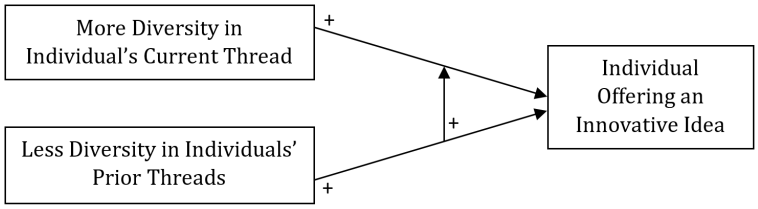

Figure 3.5.1: Evolutionary Model of the Diversity Effect on Innovativeness

diverse discussion threads. We argue that the initial less diverse discussion threads provides the individual with a creative capacity which then helps the individual to leverage the later diverse discussion thread to generate new ideas.

3.5.1 Implications for Research

By distinguishing the two time periods of diversity (prior and current), we are able to integrate the conflicting research findings to explain why diversity may have both positive and negative effects on innovation in crowdsourcing. Moreover, these two different time periods of diversity provide a mechanism for incorporating task interdependency across time (Puranam et al. 2013), by firstly accounting for less diversity among contributions in earlier discussion threads compared to more diversity among contributions in later discussion threads. This temporal orientation or flow at the individual level suggests that temporality and time is essential in analyzing diversity in crowdsourcing, although it is usually not assessed (Ancona et al. 2001). Thus, we refer to our findings as an evolutionary model of diversity for innovation in crowdsourcing. This model is graphically depicted in Figure 3.5.1.

Research in crowdsourcing diversity has focused on diversity of ideas generated by different individuals (Bayus 2013; Boudreau 2012; Brabham 2008; Frey et al. 2011a). Some crowdsourcing platforms are even designed to foster a set of diverse ideas that are later forced to be integrated or converge to generate innovative ideas (Malhotra and Majchrzak 2014). Yet, aside from research on crowding behavior (Boudreau 2012; Hutter et al. 2011), prior research on crowdsourcing has essentially ignored the initial conditions of diversity in which individuals enter the crowdsourced discussion, and how changes in those conditions may affect innovative behavior. In our evolutionary model, we suggest that, for innovation to eventually emerge from the crowdsourced discussion, individuals must be able to enter the crowdsourced discussion in a manner that enhances their shared agreement (Carlile and Rebentisch 2003), and then proceed at their pace to

a more challenging discussion thread to be stimulated for innovation (Malhotra and Majchrzak 2014; Page 2008; Tsoukas 2009).

This evolutionary model captures the advantages and disadvantages of diversity in a comprehensive way in crowdsourcing by considering time. Our findings suggest that individuals should enter a crowdsourced discussion that appears to have participants similar in the level of expertise, and then jumping over to a different discussion thread with more diverse individuals. Our findings suggest individuals who select discussion threads in the reverse order will be less innovative. Similarly, our findings suggest that individuals who only engage in diverse threads, or only engage in less diverse threads will also not be innovative.

Finally, the evolutionary model in crowdsourcing may shed light on why there is a loss of innovativeness in brainstorming groups compared to the same individuals working separately (Paulus and Nijstad 2003). Our model suggests that individuals who have shared agreement might better cope with the evaluation apprehension from other participants – an apprehension that is associated with a decrease in a productivity loss of ideas (Claxton 2000). This suggests that organizational researchers should consider an intervention to build shared agreement among participants. Moreover, this online setting allows controlling for free-riding because individuals are posting alone and production blocking (i.e. an individual who cannot produce since the communication medium is being used by another) problems, compared to other causes of productivity loss in brainstorming (Paulus and Nijstad 2003). In contrast to results in physical world where they have a constant interaction that can build trust, online teams need to first build this trust through participation and later exploit it.

3.5.2 Implications for Practice

By considering the different effects of diversity of current and prior threads, we are able to suggest how an individual and the organization sponsoring the crowdsourced discussion can effectively use diversity based on prior diversity. This allows organizations to automatically generate a set of tools based on prior individual background to effectively manage diversity since it does not require a qualitative content analysis that is not feasible in real-time. For organizational designers and practitioners, the evolutionary model suggests that, at the beginning of a crowdsourced discussion, the crowdsourced discussion should be populated with a wide range of discussion threads – either intentionally or serendipitously. The designers should encourage participants join homogeneous discussion threads by showing discussion threads that are homogenous (i.e. to

ensure low prior diversity). Then, part way through the crowdsourced discussion, the participants should be encouraged to try alternative threads with greater diversity (e.g. via personalized emails or newsletters). Individuals can recognize easily (e.g. posts with different perspectives vs. posts with similar perspectives) when diversity among the participants exist, and thus can self-select into new discussion threads that fulfill this requirement. As such, crowdsourcing to enhance the likelihood of obtaining strategically useful and innovative ideas increasingly becomes a more controllable process than what is currently done today.

3.5.3 Limitations and Future Research

Although this analysis was conducted across multiple crowdsourced discussions, there may be limitations in extending these results to firms sponsoring other crowdsourced discussions which are non-governmental, non-strategic and competitive. The present study is limited to posting data with no access to what posts were viewed. The novelty of these posts was assessed following the definition of Amabile (1988) which identifies novel and implementable ideas for the organization by the panel of experts. Finally, the measures of diversity were self-report, raising the question as to whether more objective measures of diversity would yield the same results.

Our evolutionary model raises several questions worthy of further investigation. Since not all individuals follow the proper evolutionary path, what types of individuals are more likely to select the proper path? Do creative individuals recognize the value of this evolutionary path and follow it intentionally or does the evolutionary path promote innovation even among non-creative individuals? Does this evolutionary path only apply to crowdsourcing where it can be more easily monitored, or does it apply to the design of teams as well? Is there an agency issue here such that it is only those individuals who naturally select the right evolutionary path that are able to use the evolution to stimulate their innovativeness? Or can individuals be encouraged by organizational design and inducement to engage in this evolutionary path even when it differs from what they initially intended to do? Does individuals' output produce a diverse content according to this evolutionary path during the crowdsourcing?

3.6 Conclusions

Organizations are increasingly interested in successfully engaging crowds in collaborative and critical strategic decisions by tapping into the crowd's diversity. Our findings indicate that organizations should seek out, or grow, individuals with creative capacity by participating in low

diverse discussions early in the crowdsourced discussion, and later join high diverse threads to stimulate more innovative ideas. This paper argues for an evolutionary model of diversity. Finally, organizations can easily foster innovation by mechanisms that match these individuals with the appropriate amount of diversity at a time.

	Control Model	Full Model
Fixed Effects:		
Intercept	-3.96^{***}	-3.26^{***}
	(.83)	(.94)
Expertise of the individual	$.18^{**}$	$.18^{*}$
	(.07)	(.07)
Average expertise of the focal thread	−.03	.16
	(.16)	(.20)
Average expertise of prior threads	.28	−.10
	(.27)	(.31)
Engagement of the individual	$-.20^{*}$	$-.18^{*}$
	(.07)	(.08)
New knowledge produced	$.22^{**}$	$.19^{**}$
	(.07)	(.07)
Cross-poster	.18	.18
	(.15)	(.16)
H1: More current diversity		$.31^{*}$
		(.13)
H2: Less prior diversity		$.36^{***}$
		(.11)
H3: More current diversity x Less prior diversity		$.17^{*}$
		(.08)
Random Effects:		
Variance of individuals' intercept	0.47^{***}	0.48^{***}
	(0.69)	(0.69)
Variance of crowdsourced discussion's intercept	$.62^{***}$	$.67^{***}$
	(.79)	(.82)
Variance of thread within crowdsourced discussion's intercept	1.09^{***}	1.15^{***}
	(1.04)	(1.07)
Parameters:		
Conditional R^2	.41	.44
Degrees of freedom	10	13
AIC	2224.0	2218.3
log *Likelihood*	−1102.0	−1096.1
Deviance	2204.0	2192.3
ϕ^2		11.8^{**}

Table 3.4.2: Crossed Multilevel Logistic Regression Models

There occurs the beautiful feeling that only humanity together is the true human being, and that the individual can be cheerful and happy only if he has the courage to feel himself in the Whole.

Johann Wolfgang von Goethe

4

Tapping the Innovative Business Potential of Innovation Contests[1]

4.1 Innovation Contest

An Innovation Contest[2] describes an approach used by firms that encourages a public crowd to co-creatively develop innovative responses to a firm's question (Füller et al. 2014; Hutter et al. 2011). The question prompt is often quite open, such as asking the crowd to offer recommendations for new business models, new sources of revenue, or new strategic priorities (Majchrzak and Malhotra 2013). Innovation Contests or Open Innovation Challenges are derived from the Open Innovation paradigm which "assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as the firms look to advance their technology" (Chesbrough 2006).

[1] A version of this chapter is published in Business Horizons by Armisen A. & Majchrzak A. (Volume 58, issue 4, pages 389-399, 2015) a journal with a Normalized Impact per Paper of 1.008

[2] Innovation contest is used as a form to name collaborative based crowdsourcing

Innovation Contests use what are referred to as 'web-based crowdsourcing platforms' in which a firm's question is outsourced to an undefined group of participants (Estellés-Arolas and González-Ladrón-de Guevara 2012; Howe 2006). Predictions are that by 2017, crowdsourcing will be used by more than 60% of firms as a way of engaging external parties in making a wide variety of decisions with a wide variety of providers (McIntyre 2013). Innovation Contests ask the public to not simply contribute ideas but to collaborate with each other online to co-create innovative answers to the question prompt. Because many firms are dissatisfied with the incremental nature of the ideas suggested from traditional crowdsourcing, the extension offered by Innovation Contests is increasingly used by firms today (Majchrzak and Malhotra 2013; West and Bogers 2014).

This paper explores how to tap the potential of innovation contests by improving the discussion's variety, fostering participants to add their perspectives, and encouraging first time posters. There exist some methods, techniques and tools to foster creativity in non-online groups (e.g. De Bono (1985)). Theory of De Bono (1985) is based on the fact that the human brain thinks in distinct ways, represented with hats of different colors (i.e. managing, information, emotions, discernment, optimistic response and creativity), seeking to avoid ordinary 'danger-less' thinking. Bono's Six Thinking Hats process seeks to introduce parallel thinking after an objective is defined using the Blue hat that avoids adversarial confrontations by participants wearing different hats. A clear analogy to the case of innovation contests can be drawn since the organization defines the problem (Blue hat), and participants can contribute by adding their perspectives. Finally, this paper's guidelines and Bono's six thinking hats converge into maximizing the amount of different perspectives in a discussion that leads to posting of more innovative ideas.

Guidelines for managing crowdsourcing abound, including the design of the innovation contests web platforms, properly aligning rewards with motivations (i.e. participants in Innovation contests are not only motivated by winning the announced prize but also are motivated by a range of other factors such as learning and social support (Hutter et al. 2011), use of non-technical language, defining objectives, clarifying terms and conditions for participants, clearly communicating the idea selection process, and finding internal champions to implement crowd ideas (Alexy et al. 2011; Boudreau and Lakhani 2009; West 2009). These guidelines are important for establishing an environment that encourages participation. However, these guidelines generally ignore the person's context of the online discussion. These guidelines ignore the possibility that the manner in which participants behave during the innovation contest may affect whether they offer innovative recommendations or contribute to other participants offering innovative recommendations during the contest. Thus, previous research has treated the process of the

contest as a 'black box'. A first step in understanding this black box is characterizing elements of the online discussion context that may affect a person's propensity to post an innovative idea.

The online context in the innovation literature is a key factor affecting the innovativeness of the contributions (Füller et al. 2014; Hutter et al. 2011; Majchrzak and Malhotra 2013). There are three different elements that we use to characterize a person's online context: 1) amount of variety among those contributing to the discussion prior to the person's innovative post, 2) amount of collaborative vs argumentative posts that have been made in the discussion prior to the person's innovative post, and 3) whether the discussion includes previous posts from the person prior to the innovative post. In this article, we describe recent research to develop guidelines specifically directed at understanding three elements to characterize an innovative person's contribution based on his/her context. We use a quantitative analysis of a case study of an Innovation Contest to form the basis for these guidelines.

4.2 Background on the Case of an Innovation Contest

The United States Agency for International Development (USAID) is a U.S. federal agency tasked to provide funding and expertise globally to end extreme global poverty and enable resilient, democratic societies to realize their potential. Typically, USAID develops 5-year strategic missions based on internal management discussions. In this case, for the first time, USAID asked citizens in developing countries with past USAID presence for their views on which "grand challenges" USAID should address in the forthcoming 5-year strategic plan. The USAID case is an Innovation JAM (Bjelland and Wood 2008). Similar to an innovation contest but without prizes (Bjelland and Wood 2008; Hutter et al. 2011),jams encourage collaboration-based crowdsourcing (Afuah and Tucci 2012). The USAID Grand Challenges Innovation Contest was announced using current and past USAID mission staff and the extensive contact network USAID had established around the world. Social media channels, distribution of flyers, videos, pictures, and blog messages were part of the promotion plan. USAID partnered with several organizations to help recruit participants. To expose as many people from around the world to the upcoming Innovation Contest, USAID asked other organizations to announce it including non-government organizations, embassies, government agencies in countries in which it had missions, and educational institutions with an interest in global development. Figure 4.2.1 shows the front website for the Contest. The Contest ran for seventy-two hours. 254 individuals from 49 countries participated, generating 591 posts. Example posts are shown in Table 4.2.1.

Level	Subject	Body
Thread	The advancing of Entrepreneurship starts small	Dr. Paul Chandler wrote a book: "Bound to the Hearth by the Shortest of Tether" about his experiences in rural China, Brazil, and other areas where there is a "lower class". Most of these places have no money, but given the right opportunities to better themselves can become small success stories. What he expresses is the need for people all over the world to gain human rights equivalent to United States Constitution's 5th Amendment Takings Clause. This is the Use of property USUFRUCT (right to use property as seen fit); Exclusivity (right to keep others off one's land); Transferability (right to buy or sell land or pass on to children); Permanence (Has Usufruct, exclusivity and transferability forever). All of these combine to create an atmosphere of Stewardship and a good steward will make his or her property better, worth more, and more useful because of private ownership. The steward has a permanent stake in the land use and will use it to the best of his/her ability. Bad stewards will sell their land to good steward or learn how to take care of their lands better and make them more productive. This self-interest and risk will create discipline to work for the best quality and use which in turn leads to good managers being copied and making the cycle spiral upward. The role of governments in this? Enforce laws that protect private property and keep politics out of private affairs. In all things we have to start at the bottom and work our way up, trickle-down economics works in theory but as everyone knows is lousy in practice. It actually make it more worthwhile to not pay works and lay them off rather than give them a competitive salary where they can even afford the products they are helping to manufacture.[...]
Comment	Property Rights	Hi Kristopher, Your post assumes that the poor actually own land. Only 1% of the world's land is owned by women, whereas 70% of the poor are women. I hope you see the connection here. Furthermore, the poor that do own land or have settled on government land face incredibly high barriers obtaining land titles. This often prevents them from utilizing the land for enterprise for fear it will be taken away from them due to the lack of a piece of paper. See Hernando de Soto's research and efforts on this issue.
Reply	Thanks for the insight	I had no idea about those statistics, so I have to say I must step back to those with better information, however, some of what you said does prove a point on my end as well. People do settle on land but fear being forced off of it, which could in turn lead to abuses of the land because technically they feel no responsibility for it, no stewardship. Perhaps the problems lie in the laws that prevent people from having official ownership of land, even if they have lived on if for a significant period of time. I know that in the south western US there are some laws still on the books that allow for people to gain ownership of land as long as no one else has claim to it and they make improvements on it over a period of two years or so. This will not be an easy issue to work on but certainly it is important, especially to increase ownership of land to women. Thank you for reading and replying, Kristopher

Table 4.2.1: Example discussion that has thread, comment and reply level

Figure 4.2.1: Innovation Contest from Global Pulse 2010.

To assess the "innovativeness" of a post, we used criteria developed by USAID. After the completion of the Contest, a jury of staff members in the strategy unit at USAID read the 591 posts to see what they could glean from the tenor and nature of the posts; they did not engage in any systematic coding or analysis of the posts. The posts stimulated the staff members to develop what they referred to as four "innovative" (i.e., not previously tried by USAID) strategic thrusts for USAID (Ferguson 2010). The four are shown in the left column of Table 4.2.2, with definitions in the next column.

To determine whether an individual post was innovative, we followed a procedure by Lamastra (2009) in which the USAID jury's list of innovative strategic thrusts (Ferguson 2010) was used to train two research assistants about the content considered by USAID to be innovative; the two assistants then read each of the 591 posts to assess whether (no/yes) the content in a post replicated the content in the USAID list of innovative strategic thrusts. The two research assistants first independently read and coded 10 posts then met to agree on their coding, then independently read and coded the next 200 posts followed by a meeting to discuss and resolve differences, then independently read and coded the remaining 380 posts followed by a discussion to resolve differences. A Kappa inter-rater reliability coefficient was calculated for the pre-discussion codes and found to have a moderate agreement[3] (Rietveld and Van Hout 1993). Examples of posts

[3] A kappa of 0.81

Table 4.2.2: Strategic Thrusts and Example Coded Post

Innovative Strategic Thrusts (by USAID)	USAID description of thrust	Example post coded as matching the strategic thrust
Build from the experience of complex science and technology to identify, describe, and prioritize challenges, scan for patterns and systems failures, and determine how to address the challenges	Employ a systems thinking approach, take into account the inter-relationships between challenges and activities addressing them. Rather than assuming linearity, work on multiple, concurrent, parallel, or related tracks	We do need to see the large scale effectively, but I do not believe that is in opposition to small scale awareness. Large scale issues by definition deal with complex systems, and there is much talk of "scale-free" behavior in that world. (Behavior that manifests similarly at different sizes.) **I'm a big proponent of finding the patterns and metaphors in large issues, so that they are real to a person's experience in the moment.** [...]
Employ Web 2.0 technology to identify issues, converse, generate ideas, and co-create solutions and change	Imagine a Facebook-like platform allowing development professionals, farmers, medical field workers, university staff, corporate executives, and indigenous rural women to engage in dialogue and solution-generation together, work with think tanks, learning circles, digital or interactive media	Some years back the then-President of Nokia had a vision for the company. **"Internet in every pocket"**. What if we made that a Grand Challenge for **all people across our planet, literally?** And then follow that up with focused efforts to make available through that ubiquitous access, content and linkages that can provide locally **accessible educational, health services, agricultural pricing and practice information along with links to potential markets-buyers , business related content, etc., etc.** [..]
Create a "UN Agency for Global Business Facilitation", a partnership between global government and businesses	The goal of the agency would be to foster more conducive environments for doing business; facilitating start-up, operation, and management of businesses, including procedures for import and export and compliance on standards, safety, and environmental conditions	**The greatest challenge facing humanity is that of building global governance to confront all the many challenges inherent in globalization. Global governance is NOT World Government, but finding the means as an international community to bring the challenges of runaway globalization under some universal control and to harness the positive elements of globalization with a set of interlocking regimes and structures with the UN at its core.** [..]
Leverage lessons learned for from city sustainability plans	Use the city sustainability plans of San Francisco and Portland as models. Adopt a back-to-basics approach, starting with food, water, shelter, and security. Grow more food with less water; support sustainable farming; encourage a healthy lifestyle; and clean water	Many countries are suffering drought currently, like Australia and especially China lately. Consequently, CPI may increase caused by drought. People living in that country are not just **suffering lack of clean water, but also high prices of goods and services. The most impressive advertising about environment** [...]

coded as matching an innovative strategic thrust is shown in the last column of Table 4.2.2. This rigorous procedure resulted in a subsample of 64 "innovative" posts (i.e., posting matching an innovative strategic thrust), and 527 posts that were not innovative.

4.3 Elements of Online Discussion Contexts

There has been equivocality in the innovation literature about the effect of three elements of online discussion contexts on innovation. Therefore, we focused on these three different elements of person's contribution associated to their context to assess the extent to which the context affects whether an individual will offer an innovative post in an innovation contest. The three elements: 1) amount of variety among those contributing to the discussion prior to the participant's innovative post, 2) amount of collaborative vs argumentative posts that have been made in the discussion prior to the participant's innovative post, and 3) whether the discussion includes previous posts from the participant prior to the innovative post. Examining the literature on these three elements led to the following three research questions:

1. Does the variety of participants' background contributing to a discussion thread effect whether a person will post an innovative idea?

2. Do online discussions in which participants primarily add their perspectives rather than argue with others effect whether a person will post an innovative idea?

3. Do online discussions in which a person repeatedly contributes effect whether the person will post an innovative idea?

These three research questions explore different sides of the context for a poster. Note that only the third research question concerns previous behaviors of the person contributing the innovative post. The other two research questions concern the context of the online discussion's context in the thread that has been developed by participants who do not post the innovative idea for that discussion thread. This focus on others in the discussion thread, not just the innovative poster alone, is unique among the research in crowdsourcing. We briefly review the literature for the three research questions.

4.3.1 Context Element #1: Variety of Those Engaging in a Discussion Thread

Some researchers have suggested that participants in a contest who have more experience and familiarity with the topic will know what content is innovative and have the capability to generate the innovative ideas (Leonard and Rayport 1997; Schulze and Hoegl 2008; Ulrich 2011). For example, Von Hippel (2007), suggests that lead users, who are most familiar with a company's products and services, are in the best position to contribute innovative ideas for new products.

However, other scholars have suggested instead that the ability of any single individual to contribute innovative ideas during crowdsourcing may be less a function of that individual's degree of familiarity, and more a function of how varied the familiarity background is of those who participate in the discussion prior to the innovative post being contributed (Armisen 2014; Frey et al. 2011b; Howe 2006). Variety among those participating in groups has been demonstrated as resulting in the sharing of sufficiently different perspectives to inform and spark new ideas in others, thereby, leading to more innovation (Maznevski 1994; Paulus and Nijstad 2003).

Consequently, we decided to examine the role of variety across different levels of familiarity. We examined variety of all those participants in a discussion thread prior to a post being contributed that was rated as innovative or not. That is, if the fourth post in a discussion was rated as innovative, we calculated the amount of differences between participants in a discussion thread prior to someone posting the innovative idea. We were able to examine variety of background because USAID had required that as participants registered for the Contest, they answered the question of how familiar they were about the topic of Grand Challenges for USAID: "Not familiar", "Just a little familiar", "Somewhat familiar", "Very familiar", or "I have worked in/presently work in this field". Participants were evenly distributed among the different familiarity levels.

4.3.2 Context Element #2: Whether Participants Focus their Contributions in the Discussion Thread on Arguing versus Contributing New Perspectives

In any Innovation Contest, an individual may make a "top-level post" which is a post that starts a discussion thread. They may also make a comment in a discussion thread, or a reply to a comment that was posted by someone else. Based on close examination of participants of discussion threads using the collaborative technology of wikis, it has been found that when participants add comments, they are generally adding their own unique perspective to a discussion (Majchrzak et al. 2012a), such as by adding a new answer to a question posed in the top-level post. Returning to Table 4.2.1, an example of adding a new perspective as a comment to a top-level post is shown.

In contrast to adding new perspectives, are posts that engage in back-and-forth comment-reply-comment-reply argumentation; research has shown that these back-and-forth replies are typically a deliberation over an issue of conflict (Kane et al. 2014; Tausczik et al. 2014), such as by arguing about whether a solution is the correct one.

From the existing literature, there appears to be two schools of thought on the effect of arguing vs adding perspectives on generating innovation. Some scholars argue that back-and-forth replies foster innovation because it encourages participants to continuously contrast their perspectives with those of others, which in turn creates a motivation to resolve the differences, which causes conceptual reframing and more innovative ideas (Tsoukas 2009). In contrast, other scholars find that arguments often devolve non-innovative interpersonal conflict and thus adding new perspectives is more likely to generate innovative solutions because others reading the new perspective will be able to note the differences without negative confrontation (Majchrzak et al. 2012a). So the question we addressed was whether participants were contributing as comments (adding) or rather as replies-to-replies prior to the innovative post.

There were 128 top-level posts ranging from threads of only 1 post to threads of 7 posts. The 591 posts were distributed across the 128 threads, with some posts as top-level posts, others as comments, and others as replies. The subsample of 64 innovative posts was also distributed such that some were top-level posts, some were comments, and others were replies.

4.3.3 Context Element #3: Has Person Posting Innovative Idea Previously Posted in Discussion

A distinction can be made in participation behavior between those who post once and only once, and those that post more than once. Of the 254 participants, 149 were single posters, and 105 were repeat posters. Previous research on "observers" indicates that most single posters have observed for substantial periods of time before posting (Faraj et al. 2011). Therefore, single posting behaviors are indicative of a contribution purposefully inserted after observing others' behaviors in the Contest, followed by observations of the effect of that contribution on continued discussions.

Scholars are of two orientations regarding the effect of previous posting on innovativeness of a future post. Some scholars argue that the most innovative ideas come from those posted by "peripheral" members (Kane et al. 2014) who only post once. These peripheral members do not feel a part of the community of other posters and are thus not hindered by social norms about

what would be considered appropriate ideas to post. Because their posts are not bound by existing social constraints, these peripheral members are thought to offer more innovative posts.

In contrast, other scholars argue that the most innovative ideas come from repeat or core participants because they are the ones who have spent the most time contrasting their perspectives with others, are most likely to be brokers between different perspectives, and are most motivated to resolve contrasting perspectives and cognitively reframe their views to generate an innovative post (Feller et al. 2012).

Therefore, we examined whether innovative posts were being offered by those who only posted once versus those who posted repeatedly.

4.4 Findings

Our analysis first focused on each of the three research questions individually. The relationship between the three research questions requires a larger sample size than we have, and thus is relegated to suggestions for future research. We found three different ideal profiles for generating innovative posts.

4.4.1 Profile #1: Discussion Threads with Varied Person Backgrounds

To examine the effect of contributors' variety of discussion threads on innovative posts, we used the Teachman formula for variety across the familiarity of participants engaged in the discussion prior to each post (Harrison and Klein 2007). We calculated variety for each discussion thread prior to an innovative post, and each discussion thread prior to a non-innovative post (Teachman 1980).

We found a significant difference between discussion threads prior to innovative vs non-innovative posts. That is, variety in the discussion thread is higher when prior to an innovative post (mean=0.90) than when prior to a non-innovative post ($mean = 0.77$). This was a significant difference: $t(54) = -2.43$, $p = .02$, as shown in Figure 4.4.1. A robustness check using a different form of variety based on the different types of jobs (Executive, Small business owner or Entrepreneur, Project or Program Manager, Staff, Student, Consultant, Other, Teacher or Educator, Unemployed, or Volunteer) was used. It was also significantly higher ($t(54) = -2.63$, $p = .01$) when prior to an innovative post ($mean = 1.32$) than when prior to a non-innovative post ($mean = 1.11$). Moreover, after controlling for level of familiarity of the individual making the innovative post, the difference was still significant. This suggests that variety among those

earlier in a discussion thread affects whether an innovative post is contributed later in that discussion thread. This suggests that the more varied the participants in a discussion thread, the more different perspectives they share, increasing the exposure of other participants to different perspectives which sparks an innovative thought that is contributed.

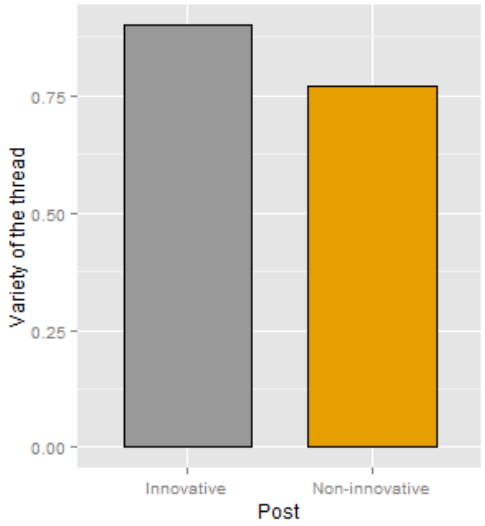

Figure 4.4.1: Significant difference in the average diversity in the current thread where the innovative post occurred is more likely to be associated with innovation ($M = 0.90$) and non-innovative ($M = 0.77$), conditions; $t(54) = -2.41$, $p = .02$.

4.4.2 Profile #2: Discussions Threads Where Contributions Add Perspective Not Argue

We calculated the position of the post, either being a top-level, comment-level or reply-level in the discussion. If arguing fosters innovation, then we would see more reply-to-replies occurring in the discussion thread prior to the innovative post. If adding new content fosters innovation, then we would see more comments occurring in the discussion thread prior to the innovative post.

As shown in Figure 4.4.2, we found that position of the posts (top level, comment-level or reply-level) were more likely to be top-level or comment-level discussion threads when prior to an innovative post ($mean = 0.55$) than prior to a non-innovative post (0.38). This was a significant difference $t(50) = -2.11, p = .04$, indicating that, on average, innovative posts were more likely

to occur in discussion threads composed of comments – construed as adding perspectives – than reply-to-replies – construed as argumentative. By offering comments instead of reply-to-replies, participants are avoiding direct conflict and a narrowing of the discussion to the point of conflict among two people. By offering more and more comments, more and more content and different perspectives are shared, helping participants to spark new innovative ideas.

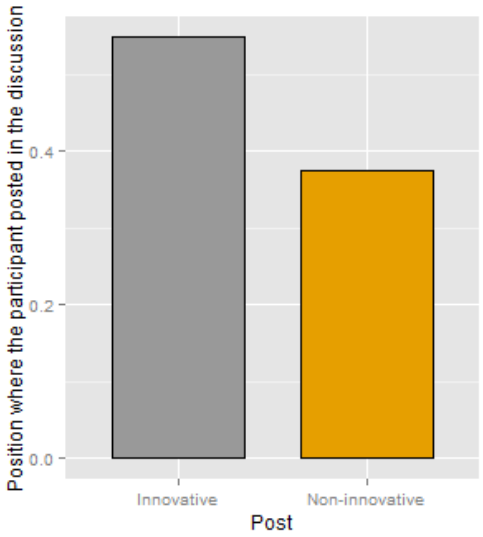

Figure 4.4.2: Significant difference in the position where the person posted in the discussion (3.. top level, 2.. comment level, 3.. reply level) with innovative ($M = 1.63$) and non-innovative ($M = 1.45$) posts $t(51) = -2.11$, $p = .04$.

4.4.3 Profile #3: First Time Posters

We subdivided the sample of all participants into those who posted only once – referred to as "first time posters" – and those who posted more than once. Of the 105 posting more than once, the median number of posts was 3. We conducted a chi-square test to determine if repeat or single posters were more likely to offer innovative posts. As shown in Figure 4.4.3, the percentage of innovative posts was significantly higher among the first time posters (15%), than repeat posters (9%), $\chi^2(1, N = 591) = 3.77$,$p = .05$.

While first-time posters only posted once, research on posters indicate that they often spend considerable time observing the behaviors of others, prior to posting (Füller et al. 2014). By focusing on observing other participants and how their ideas could affect the community, they

focus more on adding their perspectives instead of engaging with the community for support, or help (Füller et al. 2014). Thus, first time posters bring value to the innovation contest by adding their perspective, even though they do not engage in dialogue with participants.

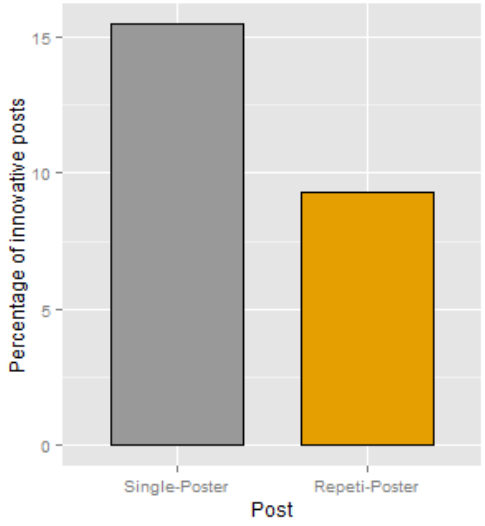

Figure 4.4.3: Percentage of innovative posts was higher among the first time posters, than the repetitive $\chi^2(1, N = 591) = 3.77$,$p = .05$.

4.5 Guidelines for Managing Crowds Participating in Innovation Contests

Tapping the innovative potential in an innovation contest can be achieved by using three guidelines developed from our findings, along with a proper reward structure that aligns the person motivation for competition and co-operation (Hutter et al. 2011). We have organized the three guidelines into a 2x2 framework shown in Table 4.5.1. The framework distinguishes between two dimensions of participation: whether the posts come from repeat vs single posters, and the nature of the posting in the discussion threads prior to the innovative post.

4.5.1 Guideline #1: Ensure Discussions Include Non-Experts

Our findings suggest that variety of person expertise with the topic – not simply expertise level - encourages innovative posts. Since most Innovation Contest are moderated by representatives

Characteristics of Discussion Thread	Person Posting	
	Repetitive Posters	First Time Posters
Varied levels of expertise represented in thread	Managers should seek to recommend through newsletter or email those threads that maximize the person potential based on his familiarity level compared to those already present in the thread.	First post of the person matters. Manager should seek to attract as many participants as possible since the first contribution of each person tends to be the most valuable.
Focus posting on adding new perspectives rather than arguing	Managers should actively avoid deep arguments and promote exchanges of perspectives at the top-level and comment-level post.	

Table 4.5.1: Framework for Participation Guidelines

of the sponsoring company, these findings suggest that moderators should encourage expertise variety, rather than just focusing on capturing expert participants that are likely to offer the same view (Von Hippel 2007). This guideline may be particularly important for Innovation Contest rather than more generic crowdsourcing (Brabham 2008) since, in such contests, the interactions among varied individuals is what drives the innovation, not simply the level of expertise. To ensure that discussions include non-experts, moderators and managers should require the crowd to register their level of familiarity with the topic; if registrations indicate only a narrow band of people participating, the Contest may need to be promoted in a way that others believe they have something to contribute. Once a range of levels of expertise are included in the registered pool, the discussion threads should be monitored to assess how the different levels of expertise are distributed within each discussion thread. If there appears to be only a narrow band of expertise contributing to a discussion thread, registered participants can be targeted and individually solicited to participate in particular discussions.

4.5.2 Guideline #2: Encourage Participants to Add Perspectives Rather Than Argue

Innovation Contests seek to solve specific problems. This research suggests that innovation is more likely when comments are made during discussion threads for adding new solutions and perspectives, rather than entering into reply-to-reply arguments. Reply-to-reply posting appears to replicate knowledge rather than adding new knowledge (Majchrzak et al. 2012b). Since arguments are rarely "won" in these contexts, except through attrition (Kane et al. 2014), reply-to-reply posting appears to waste valuable person time. Moderators, then, should monitor the progress of the Innovation Contests to note when reply-to-reply posting begins to occur in

a discussion thread, and encourage participants to offer their own perspectives as comments. In addition, incentives should be considered (Füller et al. 2010) which encourage participants to post comments as new perspectives. Such incentives can take the form of "badges" that can be turned in later for valued rewards or points that add to one's reputation as a helpful collaborator (Füller et al. 2010).

4.5.3 Guideline #3: Encourage First Time Posting as well as Repeated Posting

Our findings indicate that first time posters are important for the innovation process. While repeat posters may not post the most innovative ideas, they also are at least indirectly important to the innovation process by ensuring continuance of activity in the Contests. Without the activity, the first time posters are unlikely to come. Thus, managers should encourage people with different motivations and expected levels of effort to participate. Messages such as: "even one post makes a difference", signals to possible participants that they do not have to take on a long-term commitment or be part of a community.

4.6 Conclusion

Innovation Contests are a critical mechanism for implementing open innovation strategies in firms today. Doing these Innovation Contests correctly will help determine if new innovations can be derived from the crowd. Heeding our three guidelines should help to improve the likelihood that contributions from the crowd will be novel for the firm and implementable.

Measure what is measurable, and make measurable what is not so.

Galileo Galilei

5

Consensus in Innovation Contests Categorization by means of Fuzzy Partitions[1]

5.1 Introduction

These requests for innovative ideas are often part of what are referred to as innovation crowdsourcing events (Leimeister et al. 2009; Majchrzak and Malhotra 2013; Surowiecki 2005). In these events, an announcement is made by a sponsoring organisation to a variety of sources (via advertisements, postings on list servers, or invitations). The announcement describes the sponsoring organisation; the problem, issue, or question that the organisation is raising; information about the problem; a request for innovative ideas; a website where contributions can be made; and a timeframe (from three days to three months) during which the event will occur. Such events are increasingly used for soliciting new research & development (R&D) ideas as part of a paradigm shift referred to as open innovation (Chesbrough 2006). Government (Administration 2 03) and

[1]A version of this chapter is published in Applied Soft Computing by Armisen A., Sanchez-Hernandez G. & Majchrzak A. (Volume 35, pages 921-930) a journal with a Impact per Paper of 2.857

non-government organisations (Red Cross 2 03), as well as private sector companies (e.g., (King and Lakhani 2011; Malhotra and Majchrzak 2014)) request ideas ranging from new business strategies (Malhotra and Majchrzak 2014) to how to provide water supplies (OpenIDEO 2 03). These innovation crowdsourcing events are known as innovation contests[2] or community-based innovation contests (Bullinger et al. 2010; Leimeister et al. 2009; Majchrzak and Malhotra 2013; Piller and Walcher 2006).

In a successful innovation initiative, be it crowdsourced or more traditional methods of customer focus groups, a large number of new ideas are generated (Diehl and Stroebe 1987; Neyer et al. 2009; Simonton 1999). In fact, "if the innovation initiative is very successful, it may generate so many ideas that the selection of the most promising ones becomes very difficult or costly" (Santos and Spann 2011). An innovation initiative requires identifying both innovative ideas and interactions (e.g. posts in an innovation contest). The first step is to select a set of categorised innovative ideas from a large number of ideas found in one or more participant interactions. The chief innovation officers (CIOs) select innovative ideas using a criteria including novelty, competitive advantage, feasibility, and fit with existing competencies (Cooper and De Brentani 1984). The second step is to identify posts based on the innovative ideas selected to properly motivate participants (Füller et al. 2012a). This identification is, in general, done by a group of decision-makers or R&D managers who categorise posts as matching the set of CIO-selected ideas (Lamastra 2009). An alternative approach is for CIOs to select the innovative ideas and posts: however, this approach requires more time from CIOs (and whose time is scarce) (Toubia 2006).

Categorising posts based on innovative ideas is inherently subjective (Piller and Walcher 2006) since "the process of decision-making, in particular of a group type, is centered on humans, coming with their inherent subjectivity, imprecision and vagueness in the articulation of opinions" (Herrera-Viedma et al. 2014, p. 5). In the case of an innovation contest, the use of confidence degrees (i.e. a parameter that captures the certainty of the decision-maker or R&D manager) enables us to capture the fuzzy nature of the process and so make it closer to reality. Despite an increase in effort by decision-makers being required when using confidence degrees, their use enables harnessing information about uncertainty from each decision-maker in order to better obtain a soft consensus. Confidence degrees are obtained by means of a set of linguistic labels. The use of these linguistic labels enables us to generate a fuzzy partition per decision-maker. Finally,

[2]Innovation contest is used as a form to name collaborative based crowdsourcing

the coincidence matrix is assessed to determine the shared memberships of each post. The soft consensus degree is then computed using a fuzzy kappa (Dou et al. 2007).

The paper is organised as follows. Section 5.2 establishes the theoretical framework regarding consensus-decision-making and fuzzy partitions. Section 5.3 defines the framework and associated methodology used to translate a crisp partition into a fuzzy one by exploiting the degrees of confidence of the decision-makers. In Section 5.4 a real-case application of the methodology presented is detailed along with research findings and results. Section 5.5 discuses the implication of selecting different parameters in our methodology. Finally, the conclusions and directions for further research are explained in Section 5.6.

5.2 Theoretical framework

In this section, firstly some theoretical background information on the consensus process is presented; secondly, an overview is provided on the use of fuzzy partitions for building a set of fuzzy indices (classically used to compare a given partition with an *a priori* structure, or just another partition) are analysed to compare the fuzzy partitions obtained. Finally, these indices are used to measure agreement and extended for more than two decision-makers.

5.2.1 Consensus process

Group decision-making (GDM) is a problem-solving activity involving groups of decision-makers that produces a final choice based on the selection among several possibilities (Edwards 1977). Consensus process is a step considered in some types of GDM technique that focuses on obtaining the maximum agreement between a set of decision-makers (Herrera-Viedma et al. 2007, 2002). In general, GDM requires the majority of individuals to approve a choice, and that the minority agrees with the final choice (Herrera-Viedma et al. 2002; Kaner 2014). Other group decision-making techniques involve voting-based methods (Ayad and Kamel 2010), or structured techniques consisting of several rounds using an anonymous summary known as Delphi (Linstone et al. 1975).

Consensus decision-making techniques seek to: 1) empower all participants in the decision process; 2) listen to other participants; 3) challenge the other perspective until reaching a consensus; 4) integrate different viewpoints from other participants; 5) provide the mechanics to work on the dynamics that are not rational (such as emotions and intuitions); and 6) allocate different times to each decision if need be (Bressen 2007). These properties are better fulfilled

when using confidence degrees as a feedback mechanism since this enables the uncertainty of the decision to be captured and used later in the discussion (Edwards 1977).

Traditional approaches to consensus decision-making that do not use confidence degrees are widely used when the subjective categorisation of information based on subjective criteria (such as novelty) is needed (Saldaña 2012). This subjectiveness is handled by an inter-rater reliability assessment of the classifications assigned to the data, where two or more decision-makers analyse the data, and make various predefined categorisations (Gwet 2014). A good categorisation that overcomes this subjectivity weakness is defined when the decision-makers have an agreement that is higher than a specific threshold (Landis and Koch 1977). Traditional approaches assume that each decision-maker has the same level of decision-making certainty, and so introducing unnecessary error into the categorising by assuming that each decision-maker is absolutely certain about the categories assigned. For example, two decision-makers seeking to categorise emotional intensity in an idea may not be absolutely certain.

Decision-making in a fuzzy environment is a decision process in which the goals, constraints, or the system under control are fuzzy by nature (Bellman and Zadeh 1970). However, consensus decision-making for medium to large datasets has not adopted the use of fuzzy logic when categorising alternatives since adoption demands an additional decision-maker effort (e.g. a post to an online discussion thread would require not only the categorisation of a post to a category, but a confidence degree for each assigned category) in a vague, complex, process that might introduce possible inconsistencies among the decision-makers (Martínez and Montero 2007). Some sets of consensus decision-making techniques such as multi-attribute utility analysis use confidence degrees to capture the different decision-maker preferences (Torrance et al. 1982). Multi-attribute utility analysis has been extended to fuzzy environments using different coefficients (Chen 2000; Chiclana et al. 2013). This paper proposes capturing and using in posterior discussion one confidence degree per decision instead of multiple confidence degrees.

Capturing one confidence degree simplifies the decision-making process by capturing the fuzzy nature of the decision, while adding a minimal additional effort for the decision-makers (e.g. a post with ten categories requires one confidence degree instead of ten confidence degrees). The use of these confidence degrees helps to overcome the uncertainty management challenge (Martínez and Montero 2007) by enhancing decision-makers with additional information to assess their own decision-maker certainty through a degree of confidence. Moreover, these confidence degrees enable capturing intuition, or a viewpoint that is worth exploring if there is disagreement, by setting a low confidence degree (Bressen 2007). Thus, the fuzzy approach improves the good

properties of consensus decision-making, while enabling normalising to reflect a decision-maker's personality with regard to confidence (i.e. overconfidence, confident, or underconfidence). This is known as a soft consensus approach (Cabrerizo et al. 2009, 2010; Herrera-Viedma et al. 2007, 2005).

5.2.2 Fuzzy partitions

Within the framework of fuzzy sets introduced by Zadeh in 1965, a fuzzy set is characterised by a membership function that associates each element with a real number in the interval $[0, 1]$. A fuzzy partition of a set is then defined by giving the degree of membership or adequacy of each individual to each considered category (classes or clusters). The formal description of a fuzzy partition was introduced by Ruspini (1969). A fuzzy partition of a given set $X = \{x_1, x_2, \ldots, x_N\}$ into K clusters C_k, $k \in \{1, \ldots, K\}$, is defined by a set of values μ_{ik}, in which each value indicates the membership degree of the alternative i to the category k and satisfies the following conditions:

$$0 \leq \mu_{ik} \leq 1, \forall i, k \text{ and } \forall k \, \exists i \text{ such that } \mu_{ik} > 0.$$

In general, a fuzzy partition is expressed by a matrix $F_X = (\mu_{ik})$, where in each row, the marginal membership degree of the alternative x_i to the category C_k is represented. In addition, when the constraint $\sum_{k=1}^{K} \mu_{ik} = 1, \ \forall i$ is satisfied, the fuzzy partition is termed a fuzzy probabilistic partition. Note that in the extreme case in which $\mu_{ik} \in \{0, 1\} \ \forall i, k$, a crisp partition is obtained. Thus, fuzzy probabilistic partitions can be considered as a generalisation of crisp partitions, in the same way than fuzzy sets are considered as a generalisation of crisp sets.

To define a fuzzy degree of compatibility or similarity between fuzzy partitions, t-norms and t-conorms (denoted with a superscript '*') have extended intersection and union operators to fuzzy sets theory (Dubois and Prade 2004). The following t-norms and t-conorms are broadly used:

- MinMax:
  - Min: $M(y_1, \ldots, y_n) = min\{y_1, \ldots, y_n\}$
  - Max: $M^*(y_1, \ldots, y_n) = max\{y_1, \ldots, y_n\}$
- Probabilistic product:
  - $\prod(y_1, \ldots, y_n) = y_1 \cdot \ldots \cdot y_n$

- $\prod^*(y_1, \ldots, y_n) = 1 - \prod_{i=1}^{n}(1 - y_i)$

- Lukasiewicz:
  - $W(y_1, \ldots, y_n) = max\{1 - n + \sum_{i=1}^{n} y_i, 0\}$
  - $W^*(y_1, \ldots, y_n) = min\{\sum_{i=1}^{n} y_i, 1\}$

5.2.3 Indices to compare crisp partitions

The most classic indices to compare two partitions are described below. The Rand index (Rand 1971), and most of the following indices, are based on pairwise comparisons of data alternatives. The original crisp version of these indices handles two crisp partition matrices (P_X and Q_X) of the same set of alternatives X. Both partitions can "categorise" data alternatives into a different number of clusters. The following sets should be defined to introduce these indices:

- S_1 is the subset of tuples $\{x, y\} \in X$ paired both in P_X and in Q_X;
- S_2 is the subset of tuples $\{x, y\} \in X$ paired in P_X but not paired in Q_X;
- S_3 is the subset of tuples $\{x, y\} \in X$ not paired in P_X but paired in Q_X;
- S_4 is the subset of tuples $\{x, y\} \in X$ neither paired in P_X nor in Q_X.

Let the following numbers be the cardinal of the previously defined sets:

$$a = |S_1|,\ b = |S_2|,\ c = |S_3|,\ d = |S_4|. \tag{5.1}$$

Note that $\{S_1, S_2, S_3, S_4\}$ is a partition of the set consisting of the subsets of two elements of X and that $a + b + c + d = |X|$. On the one hand, $S_1 \cup S_4$ are the concordant pairs between P_X and Q_X (agreements), while on the other hand $S_2 \cup S_3$ are the discordant pairs (disagreements). Bearing this in mind, the Rand index is defined as the ratio between the concordant pairs and the total number of pairs. One of the main criticisms against the Rand index is that it is not *corrected by chance*, that is to say, its expected value is not zero when comparing two random partitions (Cohen 1968). The adjusted Rand index is corrected by assuming that the number of clusters to be compared is the same in both partitions (Hubert and Arabie 1985).

Another common criticism of the Rand index is that it gives the same importance to both agreement terms a and d, thus making no difference between pairs of alternatives that are joined

or separated in both partitions. Moreover, term d may dominate the other terms a, b and c. The removal of this term d results in the so-called Jaccard coefficient (Halkidi et al. 2001). It can be interpreted as the proportion of good pairs with respect to non-neutral pairs. The absense of term d is also exhibited in the so-called Fowlkes-Mallows index (Fowlkes and Mallows 1983), Minkowski measure (Jiang et al. 2004) and the ω_τ statistic (Jain et al. 1988).

The main concern of these related indices is that they do not perform properly when the clusters to be analysed are imbalanced (i.e. the probability of one cluster vastly exceeds others) (Cohen 1968). To solve this, Cohen's kappa coefficient measuring the similarity between two partitions (i.e. P and Q) was defined as follows:

$$\omega_\kappa(P, Q) = \kappa = \frac{p_o - p_e}{1 - p_e}, \tag{5.2}$$

where p_o is the relative coincidence observed between partition P and Q and p_e is the hypothetical probability of chance coincidences, using the observed categorisations to compute the probabilities of each partition randomly choosing each category. If partitions completely match, then $\kappa = 1$. However, if the categorisations differ completely, then $\kappa = 0$. The computation of both terms p_o and p_e is detailed in Dou et al. (2007).

5.2.4 Measuring Agreement: Indices to compare fuzzy partitions

The design of suitable measures for comparing fuzzy partitions is an important topic of research in several multidisciplinary areas (Kukar 2003; Osei-Bryson 2010) such as clustering. Some of the previously introduced indices in Subsection 5.2.3 have been extended into fuzzy versions: fuzzy Rand index and others (Campello 2007) and fuzzy kappa index (Dou et al. 2007).

The extension of Cohen's kappa preserves the formula given in (5.2) for computing Cohen's kappa statistic by taking into account the fuzziness of the partition in the calculation of the proportions p_o and p_e. The computation of such terms needs the selection of a t-norm in order to aggregate the membership degrees of a certain alternative to a category:

$$\kappa_\tau = \frac{p_{o_\tau} - p_{e_\tau}}{1 - p_{e_\tau}}, \tag{5.3}$$

where τ is a t-norm (see Subsection 5.2.2) and p_{o_τ} and p_{e_τ} stand for the observed and expected agreement, respectively, computed by employing the same t-norm τ. The values of p_{o_τ} and p_{e_τ} are defined by Dou et al. (2007) (see Subsection 5.3.4).

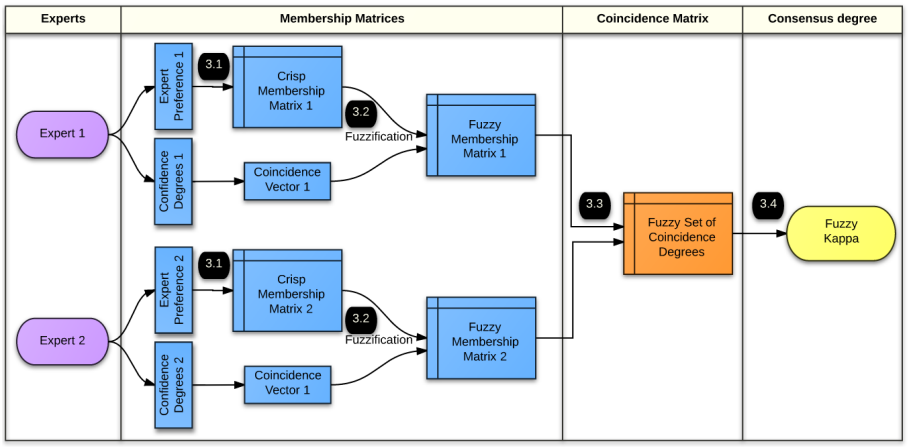

Figure 5.3.1: Scheme of the two-decision-maker soft consensus agreement proposed. Number indicates the section in which each step is detailed.

5.3 A Fuzzy Methodology to Measure Agreement in Classification

This section presents a fuzzy methodology to measure agreement based on confidence degrees and decision-maker preferences for two decision-makers. In Figure 5.3.1 the different steps of this methodology are highlighted. Initially, we present how to transform from a given crisp partition, in which each decision-maker has categorised each alternative into a unique category (i.e. "3.1" on Figure 5.3.1 and in Subsection 5.3.1), to a fuzzy partition by taking into account the confidence degree stated by the decision-maker in each classification (i.e. "3.2" on Figure 5.3.1 and in Subsection 5.3.2). This section describes the use of fuzzy indices to define a fuzzy membership matrix: the coincidence matrix between decision-makers (i.e. "3.3" on Figure 5.3.1 and in Subsection 5.3.3). This coincidence matrix is used to measure the agreement or consensus degree (i.e. "3.4" on Figure 5.3.1 and in Subsection 5.3.4). This consensus degree, as well as the decision-maker preferences and confidence degrees are used as feedback. For this reason, this framework can be classified as a feedback mechanism and consensus degree based on distance between the decision-makers according to the taxonomy given by Palomares et al. (2014). Finally, the methodology can be extended to more than two decision-makers following the steps defined in Subsection 5.3.5.

Let us assume a decision-making problem in which two decision-makers $\mathcal{E}_1$ and $\mathcal{E}_2$ analyse a set of alternatives $X = \{x_1, \ldots, x_i, \ldots, x_N\}$ to classify or assign them on a set of categories or classes $C = \{c_1, \ldots, c_k, \ldots, c_K\}$ (i.e. each decision-maker $\mathcal{E}$ states his/her selections through

his/her preferences associated to a crisp membership degree $\delta_{ik}^{\mathcal{E}} \in \{0,1\}, \forall i \in \{1,\dots N\}$ and $k \in \{1,\dots K\}$). In addition, decision-maker $\mathcal{E}$ associates with each alternative a confidence degree $D_i^{\mathcal{E}}, \forall i \in \{1,\dots N\}$, to express the level of confidence that he/she has when assigning a category to this alternative. The values of this confidence degree are given by a linguistic ordinal term set $S = \{s_0,\dots,s_g\}$ whose cardinality or granularity is $\#S = g+1$, being $D_i^{\mathcal{E}} \in S$. Therefore, an N-dimensional vector $D^{\mathcal{E}} = \{D_1^{\mathcal{E}},\dots,D_N^{\mathcal{E}}\}$ is associated with decision-maker $\mathcal{E}$.

5.3.1 Stating a crisp partition

Algorithm 1 enables us to define, for each decision-maker $\mathcal{E}_1$ and $\mathcal{E}_2$, his/her crisp membership matrix according to the initial decision-maker preferences.

1 Considering a decision-maker $\mathcal{E}$, the crisp matrix $(\delta_{i,k}^{\mathcal{E}})$ corresponding to decision-maker $\mathcal{E}$'s opinions is computed as follows:
2 **foreach** *alternative* x_i **do**
3 **foreach** *category* c_k **do**
4 **if** *decision-maker* $\mathcal{E}$ *has selected category* c_k **then**
5 $\delta_{ik}^{\mathcal{E}} = 1$;
6 **else**
7 $\delta_{ik}^{\mathcal{E}} = 0$;
8 **end**
9 **end**
10 **end**

Algorithm 1: Defining the crisp partition

Example 1. Let us assume that two averagely confident decision-makers $\mathcal{E} = \{\mathcal{E}_1, \mathcal{E}_2\}$ provide their decision-maker preference on classifying a set of four alternatives $X = \{x_1, x_2, x_3, x_4\}$ in three categories $\mathcal{C} = \{A, B, C\}$. They use the linguistic term set $S =$ *{Strongly uncertain(0), Uncertain(1), Partially uncertain(2), Neither certain nor uncertain(3), Partially certain(4), Certain(5), Strongly certain(6)}* to express their confidence, as depicted in Table 5.3.1.

In this first step we transform this data to a crisp partition obtaining the results detailed in Table 5.3.2.

	Decision-maker #1		Decision-maker #2	
	Preference	**Confidence Degree**	**Preference**	**Confidence Degree**
x_1	A	3	A	5
x_2	C	0	C	0
x_3	C	2	B	5
x_4	A	1	C	2

Table 5.3.1: Example for decision-maker preferences and confidence degrees for three alternatives

	DM #1				DM #2		
	A	B	C		A	B	C
x_1	1	0	0	x_1	1	0	0
x_2	0	0	1	x_2	0	0	1
x_3	0	0	1	x_3	0	1	0
x_4	1	0	0	x_3	0	0	1

Table 5.3.2: Transformation example for crisp partitions for the two decision-makers, four alternatives, and three categories

5.3.2 Fuzzification

To make use of the extra information provided by the N-dimensional vector $D^{\mathcal{E}}$ given by decision-maker $\mathcal{E}$, this subsection introduces a methodology to translate the crisp partition defined by $(\delta_{ik}^{\mathcal{E}})$ to a fuzzy one defined by $(\mu_{ik}^{\mathcal{E}})$. This new fuzzy partition is defined from the confidence degree vector $D^{\mathcal{E}}$, and the crisp membership values, $(\delta_{ik}^{\mathcal{E}})$, and satisfies the following conditions:

1. $\mu_{ik}^{\mathcal{E}} \in [0,1], \forall i,k$;
2. $\sum_{k=1}^{K} \mu_{ik}^{\mathcal{E}} = 1, \forall i$;
3. $\mu_{ik}^{\mathcal{E}} = 1 \iff \delta_{ik}^{\mathcal{E}} = 1$ and $D_i^{\mathcal{E}} = max(S) = s_g$;
4. $\mu_{ik}^{\mathcal{E}} = \frac{1}{K} \iff D_i^{\mathcal{E}} = min(S) = s_0$;
5. An interpolation function f is considered to define μ_{ik} when $D_i^{\mathcal{E}} \neq s_0$ and $D_i^{\mathcal{E}} \neq s_g$.

Three different functions (i.e. linear, logarithmic, and exponential) have been considered to obtain the interpolation in condition #5. Figure 5.3.2 displays on the ordinates axis the new membership degree of the category originally chosen for the alternative, according to the confidence degree shown in the abscesses axis.

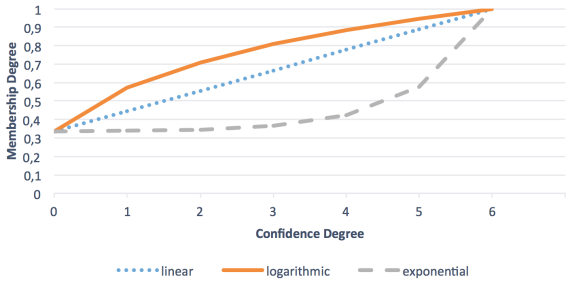

Figure 5.3.2: Example of the fuzzy membership degree of a category that had a crisp membership degree of 1. Logarithmic, linear and exponential functions are depicted by taking $K = 3$ and $\#S = 7$.

The fuzzification is a two-step process. The first step is to compute the maximum value for each alternative using one of the three interpolation functions introduced. Given an alternative x_i $i \in \{1, \ldots, N\}$ and that $k^* \in \{1, \ldots, K\}$ such that $\delta_{ik^*\varepsilon} = 1$. Then, μ_{ik^*} is defined according to Table 5.3.3.

	μ_{ik^*}	Decision-maker assumption
Linear	$\frac{D_i}{\#S-1} \cdot \frac{K-1}{K} + \frac{1}{K}$	Average confidence
Exponential	$\frac{log(1+D_i)}{log(\#S)} \cdot \frac{K-1}{K} + \frac{1}{K}$	Underconfident
Logarithmic	$\frac{e^{D_i}}{e^{\#S-1}} \cdot \frac{K-1}{K} + \frac{1}{K}$	Overconfident

Table 5.3.3: Three different functions of fuzzification

The second step is to compute the fuzzy membership degrees of the other categories, $\mu_{i\bar{k}^*}$. They are equally defined to satisfy condition #2 as follows:

$$\mu_{ik} = \frac{1 - \mu_{ik^*}}{K - 1}, \forall k \neq k^* \tag{5.4}$$

The indices defined in Subsection 5.2.2 are usually used to measure the compatibility between two partitions of the same set of alternatives. Applying the methodology detailed in the previous Subsection 5.3.2, any crisp partition with an associated vector of degree of confidence can be easily transformed into a fuzzy partition. The process for fuzzifying opinions made by a decision-maker $\mathcal{E}$ on selecting from a set C of K alternatives is detailed in Algorithm 2.

foreach *alternative* x_i **do**
  Fuzzify the vector δ_{i*} into μ_{i*}:
  foreach *category* c_k **do**
    if $\delta_{ik} = 1$ **then**
      Compute the new value of the chosen category μ_{ik^*}:
      - Linear:
        $\mu_{ik^*} = \frac{D_i}{\#S-1} \cdot \frac{K-1}{K} + \frac{1}{K}$;
      - Logarithmic:
        $\mu_{ik^*} = \frac{log(1+D_i)}{log(\#S)} \cdot \frac{K-1}{K} + \frac{1}{K}$;
      - Exponential:
        $\mu_{ik^*} = \frac{e^{D_i}}{e^{\#S-1}} \cdot \frac{K-1}{K} + \frac{1}{K}$;
    else
      Split the rest for the other unchosen categories $\mu_{i\bar{k}^*}$:
        $\mu_{i\bar{k}^*} = \frac{1-\mu_{ik^*}}{K-1}$;
    end
  end
end

Algorithm 2: Fuzzification

Example 2. Using the previous example where the crisp partitions were stated, their fuzzification is done based on the confidence degrees and linear interpolation (i.e. given by the normal confidence of the decision-makers).

Let us take the membership degree stated decision-maker $\mathcal{E}_1$ regarding the first alternative and first category $\delta_{11} = 1$, with confidence $D_1 = 3$. Given that $K = 3$, $\mu_{1k^*} = \frac{D_1}{\#S-1} \cdot \frac{K-1}{K} + \frac{1}{K} = \frac{3}{7} \cdot \frac{3-1}{3} + \frac{1}{3} = 0.667$. The others values $\mu_{1\bar{k}^*} = \frac{1-\mu_{1k^*}}{K-1} = \frac{0.667}{3-1} = 0.167$. Table 5.3.4 depicts the values obtained from this example.

	Decision-maker #1				Decision-maker #2		
	A	B	C		A	B	C
x_1	0.667	0.167	0.167	x_1	0.889	0.056	0.056
x_2	0.333	0.333	0.333	x_2	0.333	0.333	0.333
x_3	0.222	0.222	0.556	x_3	0.056	0.889	0.056
x_4	0.444	0.278	0.278	x_4	0.238	0.222	0.556

Table 5.3.4: Example of fuzzy partitions for two decision-makers in three alternatives

5.3.3 Coincidence matrix

We adopted the definition of fuzzy coincidence by (Herrera et al. 1997, p. 311) as a "fuzzy set defined on the set of decision-maker pairs and characterised by closeness observed among their respective opinions". For this reason, fuzzy kappa from Dou et al. (2007) is adapted to process values in a coincidence matrix in an approach similar to the "soft coincidence among preferences" (Cabrerizo et al. 2009; Herrera-Viedma et al. 2014; Kacprzyk and Fedrizzi 1988). This approach obtains the coincidence by means of a similarity measured among decision-makers (Herrera-Viedma et al. 2014). This coincidence matrix is defined according to one of the t-norms considered in Section 5.2.2. This enables us to give different coincidence meanings based on strict coincidence (i.e. min for the t-norm) or less strict coincidence (i.e. Lukasiewicz or product as a t-norm). Algorithm 3 details the computation of coincidence matrix for any t-norm.

foreach *alternative* x_i **do**
    Compute $fMatrix_{K*K}$ as follows:
        - Row k is the t-norm applied to element $\mu_{ik}^{\mathcal{E}_1}$ and $\mu_{i*}^{\mathcal{E}_2}$:
            $fMatrix_{k,*} = T(\mu_{ik}^{\mathcal{E}_1}, \mu_{i*}^{\mathcal{E}_2})$;
        - Column k is the t-norm applied to $\mu_{i*}^{\mathcal{E}_1}$ and $\mu_{ik}^{\mathcal{E}_1}$:
            $fMatrix_{*,k} = T(\mu_{i*}^{\mathcal{E}_1}, \mu_{ik}^{\mathcal{E}_2})$;
end

Algorithm 3: Computing coincidence matrix for t-norm

Example 3. Let us compute the coincidence matrix of the previous example assuming a product t-norm. Table 5.3.5 depicts the values obtained for the coincidence matrix.

	A	B	C		A	B	C
A	$0.667 \cdot 0.889$	$0.667 \cdot 0.056$	$0.667 \cdot 0.056$		A	0.593	0.037	0.037
B	$0.167 \cdot 0.889$	$0.167 \cdot 0.056$	$0.167 \cdot 0.056$	=	B	0.148	0.009	0.009
C	$0.167 \cdot 0.889$	$0.167 \cdot 0.056$	$0.167 \cdot 0.056$		C	0.148	0.009	0.009

Table 5.3.5: Example of coincidence matrix between experts $\mathcal{E}_1$ and $\mathcal{E}_2$ regarding alternative x_1

5.3.4 Soft consensus degree

Soft consensus degrees are based on the coincidence matrix (Herrera et al. 1997; Herrera-Viedma et al. 2014). Our proposed measure is based on the fuzzy kappa index introduced in Subsection 5.2.4. It adds robustness since it takes into account agreement occurring by chance. Algorithm 4 details the computation of fuzzy kappa index implemented according to the definitions given by Dou et al. (2007).

foreach *alternative* x_i **do**
  Compute coincidence matrix $fMatrix_{K*K}$
  Let $f_{o,p}$ be the sum of the elements of the diagonal of $fMatrix$:
    $f_{o,p} = \sum_{i=1}^{K} fMatrix_{ii}$;
  Compute $fMatrix2_{K*K}$ as follows:
    $fMatrix2 = fMatrix * fMatrix^t$;
  Let $f_{e,p}$ be the sum of the elements of the superior triangle of $fMatrix2$:
    $f_{e,p} = \sum_{i=1,i<j}^{K} fMatrix_{ij}$;
end
Accumulate all $f_{o,p}$ and $f_{o,e}$ and divide them between the number of alternatives:
  $p_{o\tau} = \frac{1}{M} \sum_{p=1}^{N} f_{o,p}$;
  $p_{e\tau} = \frac{1}{M} \sum_{p=1}^{N} f_{e,p}$;
Compute Cohen's kappa as:
  $\kappa_\tau = \frac{p_{o\tau} - p_{e\tau}}{1 - p_{e\tau}}$;

Algorithm 4: Computing fuzzy kappa

Example 4. To conclude with the previous example, let us compute κ_τ following Algorithm 4. The first step is to compute the partial $f_{o,p}$ and $f_{e,p}$

$f_{o,1} = 0.593 + 0.009 + 0.009 = 0.612, f_{e,1} = 0.037 \cdot 0.148 + 0.037 \cdot 0.148 + 0.009 \cdot 0.009 = 0.011$, $f_{o,2} = 0.333$, $f_{e,2} = 0.037$, $f_{o,3} = 0.241$, $f_{e,3} = 0.009$, $f_{o,4} = 0.322$ and $f_{e,4} = 0.032$. The expected probability $p_{e\tau} = \frac{0.011+0.037+0.009+0.032}{4} = 0.022$ and observed probability $p_{o\tau} = \frac{0.612+0.333+0.241+0.322}{4} = 0.377$ are then computed. Finally, fuzzy kappa or the soft consensus degree is obtained $\kappa_\tau = \frac{0.377-0.022}{1-0.022} = 0.363$.

5.3.5 Extending soft consensus degree to several decision-makers

In this subsection, a extension of the soft consensus degree from two to several decision-makers who each assign a set of alternatives into several categories by means of a fuzzy partition is proposed. Herrera-Viedma et al. (2014) identify two approaches to measure agreement for three or more decision-makers according to reference domain (i.e. decision-makers or alternatives). The first type is based on a fuzzy partition per decision-maker (Carlsson et al. 1992; Kacprzyk and Fedrizzi 1988), while the second type is based on a fuzzy partition per alternative (Alonso et al. 2013; Cabrerizo et al. 2010; Palomares et al. 2014). This approach can be applied to our proposed two-decision-maker soft consensus degree (as shown below) since the index introduced in (5.2) takes into account the different distribution of each category that Cohen's kappa normalises (Cohen 1968). Thus, we consider a decision-maker-based approach to extend the considered consensus degree described in three steps:

Step 1: Soft consensus degree between a pair of decision-makers, i and k: $cde_{ik} = \kappa_\tau$.

Step 2: Average of consensus degrees for each decision-maker: $ca_i = \sum_{k=1,k\neq i}^{R} \frac{cde_{ik}}{R-1}$, where R stands for the number of decision-makers.

Step 3: Aggregated soft consensus degree for the group of decision-makers: $cgd = \sum_{i=1}^{R} \frac{ca_i}{R}$.

Most real-life situations measuring consensus are based on the preferences of several decision-makers, instead of two decision-makers' preferences. This extension allows us to measure consensus degree by taking into account the different categories' occurrences.

5.4 Use Case in Innovation Contest

5.4.1 Problem definition

Several online communities in five topics created a set of ideas in an innovation contest. Among these ideas, innovative ideas or categories were identified from each innovation contest by chief innovation officers. From now on, the initially posted ideas are named as posts or alternatives, and innovative ideas are ideas or categories. The R&D managers or decision-makers then classified the community posts based on decision-maker preferences and the confidence degrees using our proposed methodology for two decision-makers. This process is depicted in Figure 5.4.1.

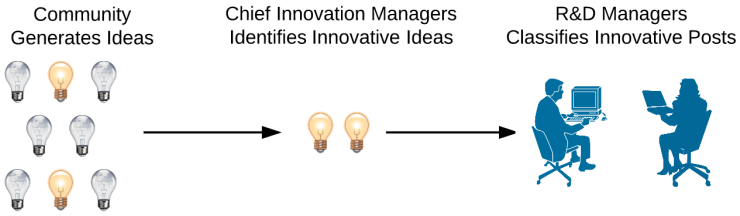

Figure 5.4.1: Use case definition

5.4.2 Data

Data was obtained from an innovation contest operated by the United States Agency for International Development (USAID), a U.S. federal agency for funding developing countries. USAID offered to the development public the opportunity to suggest development programme ideas using a 3-day innovation contest event. There were several different topics that people could focus on, including, for example, empowering women. This event was promoted by partnering with several organisations throughout the world, and 1279 individuals from 135 countries participated, generating 5268 posts across five topics.

A real-life experiment, in which the posts from five topics were categorised using the proposed methodology based on one confidence degree per post, was conducted. A jury of staff members

from the USAID strategy unit synthesised the novel ideas after reading all the posts in the challenges (Ferguson 2010). Two students per topic were then hired by the research team, trained to become decision-makers, and tasked with categorising each post into ideas, while assessing how confident they were in their decision in a scale ranging from zero to six. The first 100 posts (i.e the first batch) were used as decision-maker training and the decision-makers reached consensus after independently categorising and defining the confidence degree of each post. This set of decision-makers then independently read, categorised, and assessed their confidence degree for the second batch (i.e. a set of posts in a topic), with disagreements discussed using the confidence degree to grasp the other decision-maker's uncertainty. The process was repeated for the next batch until no batches were left to categorise and define. Table 5.4.1 provides information on how many batches, posts, and the average number of words.

Contest	# Posts per contest	Ideas	Batch	# Posts per batch	# Words per post
Empowering Women and Girls	1618	7	1	100	106
			2	200	134
			3	300	142
			4	1018	120
Pursuing Grand Challenges	591	4	1	100	215
			2	200	160
			3	291	137
Fostering Science, Technology and Innovation	768	6	1	100	119
			2	200	112
			3	468	115
Inspiring a New Generation	1515	6	1	100	136
			2	200	122
			3	300	150
			4	915	126
Building Stronger Partnerships	776	6	1	100	176
			2	200	162
			3	476	142

Table 5.4.1: Statistics of the posts and contests

5.4.3 Results

Each decision-maker assessed the idea and level of confidence for each post for a topic. Three different ways of measuring agreement were computed to assess the effect of possible decisions on the fuzzification process using product t-norm for the fuzzy kappa (see (5.3)). Moreover, crisp kappa was calculated in order to have a threshold measure. Table 5.4.2 details the results obtained.

Contest	Batch	Crisp	Linear	Logarithmic	Exponential
Building stronger partnerships	2	0.32	0.67	0.74	0.51
	3	0.26	0.79	0.82	0.68
Empowering women and girls	2	0.46	0.74	0.84	0.34
	3	0.45	0.75	0.85	0.37
	4	0.80	0.79	0.87	0.51
Fostering, science, technology, and science	2	0.23	0.82	0.86	0.67
	3	0.46	0.82	0.86	0.66
Pursuing major challenges	2	0.42	0.79	0.85	0.56
	3	0.39	0.73	0.80	0.48

Table 5.4.2: Consensus degrees using the product t-norm

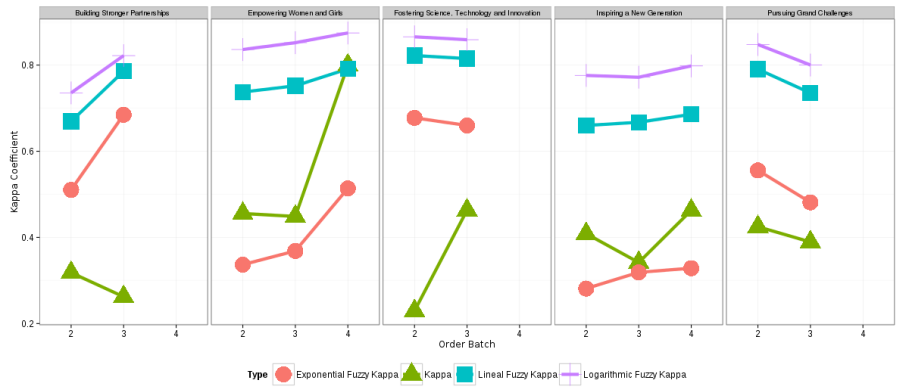

Figure 5.4.2: Agreement measurement by type of kappa and innovation contest using probabilistic product as the t-norm

Figure 5.4.2 shows that, despite using product t-norm, kappa's estimation order from each

different fuzzification normalisation function (i.e. linear, exponential, and logarithmic) is as expected. Two overconfident decision-makers show more agreement than an average underconfident decision-maker when using the same confidence degrees. An exponential fuzzification is more accurate than the traditional crisp kappa and other considered fuzzy kappa definitions (i.e. logarithmic and linear).

Consensus is clearly improving over time in all but the smaller contests (i.e. "Pursuing major challenges"). This indicates that the level of agreement improved over time as long as the data set is large enough ($>$ 600 posts). The use of confidence degrees improved the quality and efficiency of the experts' categorisation by decreasing time for each consecutive batch (as reported one expert). A possible explanation for this decrease in time is that experts do not need to be absolutely certain about their decisions since humans have their own subjectivity (Herrera-Viedma et al. 2014). This proposed method enabled us to capture this uncertainty at a reasonable cost (i.e. confidence degree for patterns instead of categories). Moreover, this confidence degree was offered as a feedback mechanism for those patterns where agreement was not reached, and so providing additional support for the discussion.

5.5 Discussion

This paper proposes a feasible method to consider the fuzzy nature of decisions when categorising a set of alternatives by means of confidence degrees. The method enables the use of different fuzzyfication functions - three of which are depicted in Table 5.3.3 - and the use of different t-norms described in Section 5.2.2. The selection of these elements has several implications that are discussed. Linear, exponential, and logarithmic functions are considered when associated respectively with normal, underconfident, and overconfident decision-makers. Overconfident decision-makers, whose subjective confidence in their judgement is reliably greater than objective judgement (Pallier et al. 2002), are compensated by using a logarithmic normalisation function during the fuzzification process. Underconfident people, whose subjective confidence in their judgement is reliably lower than the objective judgement, are compensated by using an exponential

normalisation function. Finally, people with normal confidence, whose subjective confidence is close to the objective function, are not compensated using a linear normalisation function.

The selection of the t-norm to obtain the shared membership degrees (i.e. coincidence matrix) is relevant. A minimum t-norm, that represents the standard semantics for weak conjunction logic $T(x, y) = \min(x, y)$, overemphasises total agreements among the two decision-makers, and thus reduces the relevance of disagreements. A product t-norm (i.e. $T(x, y) = x \cdot y$) penalises low partial agreement, while boosting high partial agreement. A Lukasiewicz t-norm, that is pointwise smaller than the product t-norm $T(x, y) = max\{x + y - 1, 0\}$, overpenalises low partial agreement while overboosting high partial agreement.

Several measures exist to assess the level of agreement, ranging from simple ratio agreements to measures that take into account the distribution of each category (Cohen 1968). The consensus degree defined in this paper introduces the use of a measure that takes into account the distribution of each category that is fuzzy by nature (Dou et al. 2007). Table 5.5.1 compares the typical consensus forms of measurement in three dimensions: 1) implementation complexity; 2) assumption of the distribution of categories; and 3) certainty of the decision-maker at each decision. The ratio measurement considers that all categories are equally likely, hence introducing distortion in the measurement. The kappa improves ratio measurement by considering that each category is not equally likely to happen. Finally, fuzzy kappa considers that every idea or category is not equally likely to happen and that the decision-makers are not always certain about their decisions.

Measure	**Formula**	**Complexity**	**Distribution of categories**	**Certainty decision**
Ratio	P_o^{ik}	Simple	Equally likely	Constant
kappa	$\kappa = \frac{p_o - p_e}{1 - p_e}$	Medium	Some categories are more likely than others	Constant
Fuzzy kappa	$\kappa_\tau = \frac{p_{o\tau} - p_{e\tau}}{1 - p_{e\tau}}$	Advanced	Some categories are more likely than others	Not constant

Table 5.5.1: Comparison of consensus degrees

5.6 Conclusions and Future Research

The focus of this research was to develop a fuzzy measure of agreement through fuzzy kappa based on fuzzy partitions that consider the fuzzy nature of GDM using one confidence degree for each decision (instead of for each category). This confidence degree is used to define a fuzzy partition based on decision-maker preferences. A coincidence matrix is computed by two or more fuzzy partitions using a specific t-norm. Finally, the consensus degree is obtained from this coincidence matrix.

A real use case in an innovation contest is considered. The results suggest that this methodology captures the degree of consensus more accurately, without much additional effort. Moreover, the fuzzification of the decision allows us to model and normalise different types of confidence (i.e. overconfidence, average confidence and underconfidence). Finally, the assessment of the different conceptualisations of t-norm enables us to consider less strict coincidence models (i.e. probabilistic product t-norm) or stricter models (i.e. min t-norm).

The limitations of this research are related to the constraints imposed on using real-life experiments to assess different parametrisation (i.e. fuzzification parametrisation and different types of t-norm for the fuzzy kappa). Thus, future research should estimate acceptable and optimal fuzzy consensus degrees for the different parametrisations of the kappa index.

If I had asked people what they wanted, they would have said faster horses.

Henry Ford

6 Conclusion

This chapter discusses the different chapters of the thesis (6.1), its theoretical implications (6.2), the implications to management (6.3), as well as limitations and future research (6.4) derived from the three research questions: 1) How Does the Diversity of the Crowdsourced Discussion Affect Innovativeness? 2) How do the characteristics of the crowdsourced discussion affect innovativeness? 3) How can innovativeness of ideas be rated in crowdsourced discussions to take into consideration the different levels of expertise from the decision-makers?

6.1 Discussion of the thesis

This section discusses the characteristics of crowdsourced discussions and their relation to innovativeness of the ideas produced. These findings demonstrate that the approach taken by the participant, less prior diversity (i.e. low diversity of threads from previous discussion threads), and the more current diversity (i.e. high diversity of the immediate discussion thread) have a positive relationship with innovation. This thesis tries to overcome the lack of research on how diversity in crowdsourced discussions affects innovation. Moreover, a new methodology is proposed to evaluate innovation given the fuzzy nature of innovation that is suitable for multiple decision-makers. This new methodology developed takes into account that ideas can appear in different frequency distributions in a fuzzy environment which are evaluated by multiple decision-makers. Together, the three articles empirically demonstrate that innovation in large data sets can be measured effectively, and that the characteristics of crowdsourced discussions matter for innovation.

6.1.1 How Diversity Contributes to Innovation in Crowdsourcing: an Evolutionary Diversity Model

Chapter 3 provides a model for the importance of the dimension of time to understand diversity by identifying two temporal characteristics: less prior diversity and more current diversity. Despite the diversity in discussion threads measured in different ways (i.e. expertise, job or affiliation), the results suggest that both lines of thought (i.e. the different schools of thought arguing for more and less diversity) are correct. Both lines of thought are reconciled by considering the time when assessing diversity, whether they are immediate discussion threads (i.e. current diversity) or previous discussion threads (i.e. prior diversity). Traditional crowdsourcing roles have focused mainly on assessing diversity regardless of time (Boudreau 2012; Brabham 2008; Terwiesch and Xu 2008). This chapter bridges the positive and negative effects of diversity considering time

by empirically testing the evolutionary diversity framework using an empirical analysis of six crowdsourced discussions that produced 4.628 ideas.

In this chapter, the effect of diversity is described as a positive association with more innovative crowdsourced discussions, while the positive effects of less prior diversity comes from prior discussion threads (e.g. those discussion threads in which the participant previously participated). In addition, there are spillovers effects of less prior diversity on more current diversity suggesting that participants can further exploit the diversity of current discussion threads when they build a creative capacity. This creative capacity is achieved through participation in similar discussions, which helps these individuals meet others who have different perspectives. This creative capacity can prevent participants from effectively engage cognitively in crowdsourced discussions (Paulus and Nijstad 2003).

6.1.2 Tapping the Innovative Business Potential of Innovation Contests

Chapter 4 found out that there are three characteristics in crowdsourced discussions affecting the emergence of innovative ideas into discussion threads: 1) participants from diverse backgrounds, 2) contributions adding perspectives, rather than discussion, and 3) the first contribution of the participants. The analysis was carried out using a crowdsourced discussion consisting of 461 ideas.

The first characteristic is the fact that the variety or diversity of participants in the discussion threads on innovative positions have a positive effect to innovativeness, suggesting that the most diverse participants there are in a discussion, the more diverse perspectives they share, leading to an increased innovation capacity. Although this finding is similar to that in the previous chapter, Chapter 4 focuses on the effect of the variety in the discussion thread regardless of what the participant did previously in the crowdsourced discussion. In contrast to Lakhani et al. (2007) and Terwiesch and Xu (2008), which found that the number of participants in gross stimulates innovation, the findings of this chapter are that the proportion of the types of participants matter. This finding provides additional support for Franke and Shah (2003), where the community

supports innovation by providing complementary capabilities in the group and the interaction of participants when seeking innovation.

The second characteristic is that participants generate innovative ideas when they focus on adding their perspective to the top post in the discussion thread rather than engaging in one to one discussion. On one hand, comments to the main idea help to establish multiparty discussion; on the other hand, replies to replies focus the conversation in bilateral arguments. Comments to the main idea help participants by avoiding direct confrontation with other individuals while surfacing new perspectives to the problem and generating new, innovative ideas. Blohm et al. (2011) found that participation in a collaborative environment stimulates innovation, compared to non-collaborative environments; however, it does not take into account the type of behavior that occurs within a collaborative community. Therefore, this structural characteristic expands their work by providing further evidence that innovative discussion threads seek to surface different perspective. This result also contributes to splitting feedback contribution into two components: collaboration (i.e. a comment to a main idea) and argumentative (i.e. a reply to a comment or another reply). Moreover, Chapter 4 shows that the argumentative participants who try to provide additional perspectives are more innovative than those participants who only seek confrontation.

The third characteristic is that those posts which were first posted by a participant in the crowdsourced discussion are the most innovative. These individuals are first-time posters (i.e. passive users using the taxonomy of Füller et al. (2014) or novice using the taxonomy of Kim (2000)); this is in contrast to repetitive posters who are less innovative (i.e. regulars, leaders and elders using the taxonomy of Kim (2000), or socializer, idea generator, effective contributor and master using the taxonomy of Füller et al. (2014)). This suggests that first-time posters may spend time observing the crowd (i.e. visitors using the taxonomy of Kim (2000)) before making his or her first post. This finding gives more importance to the first contribution made by participants in a crowd, which in turn it helps to explain why peripheral users are more innovative (Kane et al. 2014), because each post that an individual makes to the crowd is less innovative. A

possible explanation may be that participants become integrated and are then not able to produce innovative posts (Bayus 2013).

6.1.3 Consensus in Innovation Contests Categorization by Means of Fuzzy Partitions

Chapter 5 proposed a fuzzy measure of assessing innovation based on soft consensus decision-making when there is low agreement between decision-makers using the levels of confidence of each one. This method provided a more time-efficient approach to capture the fuzzy nature than the method based on preferences (Cabrerizo et al. 2009, 2010; Herrera-Viedma et al. 2007, 2005). Moreover, this method accounts for different frequency distances for the innovative ideas (Cohen 1968).

A fuzzy partition is a set of defined degrees of membership to specific categories, classes or clusters (i.e. ideas). These fuzzy partitions allow one to take into account the distribution of each category and the fuzzy nature of decision-making using one confidence degree for each decision. This method allows for capturing the degree of confidence of the decision-maker that the traditional Cohen's kappa method could not asses (Dou et al. 2007), using a set of confidence degrees to measure the decision-maker preferences which defines the coincidence matrix (Cabrerizo et al. 2009; Herrera-Viedma et al. 2014). This is done using specific t-norm that might be more or less restrictive depending on the setting that is selected to model. This allows us overcoming the limitations of using Kappa in overoptimistic settings (Guggenmoos-Holzmann 1996). This t-norm might be more or less restrictive depending on the setting that is selected to model.

A minimum t-norm overemphasizes total agreements among the two decision-makers, and thus reduces the relevance of disagreements. This type of t-norm is suitable for applications that focus on exploratory analysis (Kleining and Witt 2001). A product t-norm penalizes low partial agreement, while boosting high partial agreement. A Lukasiewicz t-norm over penalizes low

partial agreement while over boosting high partial agreement. This type of t-norm is suitable for applications on qualitative methods that seek to draw reliable conclusions.

Finally, each categorization made by individuals is modeled, or fuzzified, according to the individual level of confidence (i.e. under-confidence, normal, or over-confidence) (Toubia 2006). On the one hand, over-confident individuals whose subjective confidence in their judgment is reliably greater than objective judgment (Pallier et al. 2002) are modeled using a logarithmic function for the t-norm. On the other hand, under-confident individuals are modeled using an exponential function for the t-norm.

6.2 Theoretical Implications

The theoretical implications of this thesis focus mainly on crowdsourced discussions, the effects of diversity on innovativeness, and group decision-making.

6.2.1 Crowdsourced Discussions

New forms of organizations require new theories on how to leverage common ground (i.e. knowledge that is shared) during the provision of information (Puranam et al. 2013). The contributions of this thesis are about understanding the characteristics of crowdsourced discussions, which are described in the following paragraphs.

Collaboration is a dynamic process that occurs not only at the discussion thread or group level, but as an accumulation of previous interactions, as shown in Chapter 3. The reason for this could be that people are influenced by those with whom they have interacted and these effects are cumulative, as seen in Chapter 3, with less prior diversity. For example, when an individual interacting with people who think alike and generate the same kinds of questions and answers, that participant's sense of security and belonging will be increased. Then, in future interactions, the participant will be more likely to effectively engage in future discussions threads by being more open and accepting to perspectives that are different from their own.

Interestingly, the results of Chapters 2 and 3 found that diversity has a positive effect on innovation capacity when diversity is operationalized as a proportion rather than as the raw number of participants (Boudreau 2012; Brabham 2008; Frey et al. 2011a; Howe 2006). This was operationalized using a proportion measure for the current discussion thread, in contrast to how many participants collaborated in the crowdsourced discussion. Moreover, as stated in the paragraph above, participants need to build a creative capacity in order to take advantage of diversity. Therefore, this thesis suggests that the origin (i.e. whatever is prior or late) of diversity is critical because individuals in the crowdsourced discussion first must develop their creative abilities and before exploiting these abilities.

This thesis also focuses on whether a post is innovative or not by considering the path or trajectory from which the participant derived. This is achieved by assessing the prior diversity of the prior discussion threads that the individual joined. An example of this distinction relates to the relevance of the first-time poster, who is a source of innovation, as seen in chapter 4, and adding new perspectives to generate innovation (Paulus and Nijstad 2003). The relevance of the first-posts may be explained by the peripheral participants and this thesis therefore offers a complementary view to lead user literature (Franke and Shah 2003; Füller et al. 2014). This relevance is explained by the contribution of the peripheral participants that is low and innovative, yet the contribution of the lead users is high and novelty is low on average, but they have more innovative ideas. Furthermore, Bayus (2013) that found that participants' innovation capacity is diminished with the greater number of posts they make„ and Chapter 4 found that the first-post of the participants is the most innovative. These participants can decide to stay as peripheral users or seek a more active participation level and become central users.

The results suggest that, unexpectedly, the greater level of expertise one has in the field had no relation to innovation, the right combination of expertise that led to the production of innovative ideas. The innovativeness of collaboration increases when diversity comes to various sources, either from the focal discussion thread (see Chapters 3 and 5) or earlier discussion threads (see Chapter 3). Although there is a cost to increase diversity (Laursen and Salter 2006; West

and Bogers 2014) due to the prior discussion threads, diversity increases innovativeness. These results could be due to the nature of online crowdsourced discussions that aim to facilitate communication between the different indivuals (Malhotra and Majchrzak 2014).

6.2.2 Diversity's Effect on Innovativeness

The traditional literature on diversity has considered diversity as a static measure that does not change over time (Harrison and Klein 2007). However, new forms of online collaboration challenge the principle that groups are stable because in online collaboration, the groups change and evolve over short periods of time (Franke and Shah 2003). Diversity literature usually focuses on group interaction, probably due to the nature of long duration groups (Van Knippenberg and Schippers 2007). Temporary groups allow for reducing the apparently conflicting results on the effectiveness of diversity by considering time. The timing of crowdsourced discussions solves the debate on how diversity affects innovation capacity by generating an evolutionary model, as shown in Figure 6.2.1.

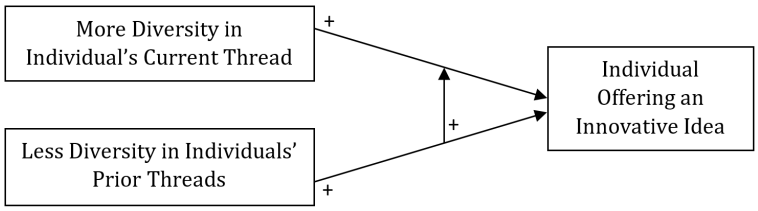

Figure 6.2.1: Evolutionary Model of the Diversity Effect on Innovativeness.

Theories of group creativity and literature (Harrison and Klein 2007) suggest that diversity leads to innovation as it allows for the integration of different perspectives (Malhotra and Majchrzak 2014; Tsoukas 2009). The results in Chapter 3 and 4 indicate that this integration can be achieved without the need to build a team, as the diversity of the extremes of the current discussion thread is associated with innovation capacity. Therefore, participants from different backgrounds produce innovative results despite how difficult it is to interact among them (Carlile 2004).

Other theories about diversity in creativity suggest that diversity hurts innovation because

individuals need individual self-efficacy for creative ability. Individuals explore whether individuals build this creative ability by evaluating the type of threads in which they participated. If these discussion threads are similar, communication could be easier. This could be because it has a higher semantic similarity between them (Carlile 2004), and thus they are more willing to accept the risks of similar individuals (Edmondson 1999). Moreover, this creative capacity could help exploit the richness of diversity in the current discussion thread or make individuals more involved in the crowdsourced discussion.

The operationalization of diversity in Chapter 3 and 4 was done using a fine-grained measure, Teachman, which allows one to capture the distribution of each category rather than the richness (Terwiesch and Xu 2008). This mode of evaluation suggests that it is not just a matter of how many people are in a crowdsourced discussion, but the proportion of the different people involved. For example, take one thousand people who work for a large organization; these people will focus primarily on the problems that are solely related to this large organization. If you instead have a community with 500 individuals from a large organization and 500 individuals from small organizations, this will allow for ideas that combine expertise with ideas of both small and large organizations, and this combination tends to be associated with innovativeness (Tsoukas 2009).

6.2.3 Group decision-making

Other implications of how to measure innovation are related to group decision-making. Chapter 5 proposes a feasible method to consider the diffuse nature of decision-making when classifying a set of alternatives through degrees of confidence rather than preferences (Herrera-Viedma et al. 2014). Levels of confidence allow one to obtain the confidence of decision-making rather than make assumptions based on the preferences of the decision maker (Herrera-Viedma et al. 2014). This method has the potential to improve the time needed to assess the subjectivity of the task (Piller and Walcher 2006) because the decision-maker does not have to expend energy to be absolutely sure about each decision. This manages to decrease the cognitive load required to assess a decision. Another important finding is extending the fuzzy kappa to more than two decision-

makers. Chapter 5 proposes to account for the similarities between pairs of decision makers that are averaged later. This approach is useful for taking into account the average agreement between a pair of decision makers; however, it does not take into account the difference in the types of categories among decision makers. Therefore, the proposed extension to more than two decision-makers is appropriate when the agreement between the various decision-makers is evaluated regardless of individual differences among the different sets of pair-wise comparisons.

6.3 Managerial Implications

Our results offer several practical implications for the management and design crowdsourced discussions. The first is that managers should seek to attract as many people as possible to join the crowdsourced discussion they are sponsoring because the first-post each participant produces is their most innovative (as explained in Chapter 4). Moreover, a manager should seek to invite a diverse group of people, and extend further invitations depending on the kind of diversity that is missing in the crowdsourced discussion in order to further enhance the innovation potential. The inclusion of a diverse set of individuals helps to produce innovative ideas (as explained in Chapters 3 and 4) and increasing this diversity should be the goal, rather than focusing on attracting as many people as possible (Terwiesch and Xu 2008).

The crowdsourced discussion should be designed to encourage discussion threads between the same types of people (i.e. having less previous diversity, as seen in Chapter 3). This is done to develop the necessary creative skills that will enable individuals to take advantage of the diversity of perspectives that can be exposed later. There are several means to achieve this, such as through the use of newsletters and personalized recommendations that can be incorporated into the crowdsourced discussion. Finally, the discussion managers may want to intervene to dissuade the argumentative discussions that occur when two individuals engage in a discussion on a one-on-one basis.

The other important implication is a new methodology to measure innovation given the fuzzy nature of decision-making. Managers should take advantage of this new methodology in assessing

the innovativeness of ideas. This methodology simplifies the process by adjusting the level of expertise of each judge (i.e. overconfidence, average confidence and under-confidence), using a parameter according to the decision taken instead of multiple parameters per each component of the decision (e.g. a level of confidence for the absence of an innovative idea), which is required when assessing the fuzzy group-decision.

6.4 Limitations and Future Research

This innovation research did not study the factors that influence innovation such as individual motivation or creativity (Bharadwaj and Menon 2000; Leiponen and Helfat 2010; Pierce and Delbecq 1977). Moreover, these results should be done in crowdsourced discussions from different sponsoring organizations where motivation surveys are obtained by surveying the participants. Finally, the methodology proposed in Chapter 5 argues that levels of confidence improve accuracy of the evaluation in decision-making; however, there is no research to date in which the acceptable threshold for claiming that a satisfactory agreement on the fuzzy kappa for different parameter fuzzification and different types of t-norm are provided. Moreover, this methodology can be adapted to obtain the same results using preferences instead of degrees of confidence as the decision-making group (Herrera-Viedma et al. 2014).

In this thesis several issues that merit further investigation have arisen related to path dependence and diversity. The first set of questions are related to path dependence, as individuals follow different paths through participation in different types of threads. These raise two questions: Is it best to participate in threads dominated by first posters? What kind of people are more likely to select lines of discussion that are less diverse in the beginning of the discussion thread and more diverse later on? Moreover, this thesis focuses on the diversity among different individuals; however, it has not been quantitatively explored whether these individuals produce different content. Therefore, new questions arise such as: Is the diversity of content also present in an evolutionary model? Is there any interaction between the individual diversity and diversity of content? Does the same result hold in an offline environment that is as dynamic as online

crowdsourced discussions? Finally, some of these results as the value of newcomers could be researched on physical teams. Finally, another research question is related to how the distance of the problem might affect the composition of the solvers. Some problems might require more experts to validate facts (e.g. defining the strategy of the firm) while other problems might instead value the opinions more (e.g. t-shirt design popularity).

6.5 Final Remarks

This thesis had three issues addressed quantitatively. Specifically, a new evolutionary framework was developed for understanding the effect of diversity on innovation capacity. In addition, participants who focused on revealing their perspectives to the group rather than engaging in one-on-one discussions were more innovative. Furthermore, those discussion threads that the individuals published for the first time in the crowdsourced discussion generate more innovative posts. Finally, a new methodology allowed for the adequate measure of innovation given the fuzzy nature of innovation and the different levels of expertise of the decision-makers.

This research focused on obtaining the diversity based on stable characteristic of the individuals rather than deriving from the content using natural language processing. Therefore, an interesting line of research would be to evaluate semantic diversity and how semantic and stable diversity are intertwined. Moreover, this thesis draws conclusions from several crowdsourced discussions to determine one organization's strategic plan for the next five years. It would be interesting to verify these results across different organizations. Finally, the proposed methodology in this thesis assesses the agreement between the different decision-makers with different levels of expertise.

References

Administration, U. G. S. 2015-02-03. "Challenge.org," http://www.challenge.gov.

Afuah, A., and Tucci, C. L. 2012. "Crowdsourcing as a solution to distant search," *Academy of Management Review* (37:3), pp. 355–375.

Alexy, O., Criscuolo, P., and Salter, A. 2011. "No soliciting: strategies for managing unsolicited innovative ideas," *California Management Review* (54:3), pp. 116–139.

Alonso, S., Pérez, I. J., Cabrerizo, F. J., and Herrera-Viedma, E. 2013. "A linguistic consensus model for web 2.0 communities," *Applied Soft Computing* (13:1), pp. 149–157.

Amabile, T. M. 1988. "A model of creativity and innovation in organizations," *Research in organizational behavior* (10:1), pp. 123–167.

Ancona, D. G., Goodman, P. S., Lawrence, B. S., and Tushman, M. L. 2001. "Time: A new research lens," *Academy of management Review* (26:4), pp. 645–663.

Armisen, A., and Majchrzak, A. 2015. "Tapping the innovative business potential of innovation contests," *Business Horizons* (58:4), pp. 389–399.

Armisen, A., Majchrzak 2014. "An Ecological Approach to Understanding How Diversity affects Collaborative Innovation during Open Strategy Development," .

Ayad, H. G., and Kamel, M. S. 2010. "On voting-based consensus of cluster ensembles," *Pattern Recognition* (43:5), pp. 1943–1953.

Bates, D., Mächler, M., Bolker, B., and Walker, S. 2014. "Fitting linear mixed-effects models using lme4," *arXiv preprint arXiv:14065823* .

Bayus, B. L. 2013. "Crowdsourcing new product ideas over time: An analysis of the Dell IdeaStorm community," *Management science* (59:1), pp. 226–244.

Bellman, R. E., and Zadeh, L. A. 1970. "Decision-making in a fuzzy environment," *Management science* (17:4), pp. B–141.

Bharadwaj, S., and Menon, A. 2000. "Making innovation happen in organizations: individual creativity mechanisms, organizational creativity mechanisms or both?" *Journal of product innovation management* (17:6), pp. 424–434.

Bjelland, O. M., and Wood, R. C. 2008. "An Inside View of IBM's' Innovation Jam'," *MIT Sloan management review* (50:1), pp. 32–40.

Blohm, I., Bretschneider, U., Leimeister, J. M., and Krcmar, H. 2011. "Does collaboration among participants lead to better ideas in IT-based idea competitions? An empirical investigation," *International Journal of Networking and Virtual Organisations* (9:2), pp. 106–122.

Boland Jr, R. J., and Tenkasi, R. V. 1995. "Perspective making and perspective taking in communities of knowing," *Organization science* (6:4), pp. 350–372.

Boudreau, K., and Lakhani, K. 2009. "How to manage outside innovation," *MIT Sloan management review* (50:4), p. 69.

Boudreau, K. J. 2012. "Let a thousand flowers bloom? An early look at large numbers of software app developers and patterns of innovation," *Organization Science* (23:5), pp. 1409–1427.

Boudreau, K. J., Lacetera, N., and Lakhani, K. R. 2011. "Incentives and problem uncertainty in innovation contests: An empirical analysis," *Management Science* (57:5), pp. 843–863.

Boudreau, K. J., and Lakhani, K. 2015. "Innovation Experiments: Researching Technical Advance, Knowledge Production and the Design of Supporting Institutions," *Forthcoming in Innovation Policy and the Economy* (16), pp. 16–001.

Brabham, D. C. 2008. "Crowdsourcing as a model for problem solving an introduction and cases," *Convergence: the international journal of research into new media technologies* (14:1), pp. 75–90.

Bressen, T. 2007. "Consensus decision making," in *The Change Handbook: The Definitive Resource on Today's Best Methods for Engaging Whole Systems (2nd. Ed.),* , San Francisco, CA: Berrett-Koehler Publishers, pp. 212–217.

Bullinger, A. C., Neyer, A.-K., Rass, M., and Moeslein, K. M. 2010. "Community-Based Innovation Contests: Where Competition Meets Cooperation," *Creativity and innovation management* (19:3), pp. 290–303.

Cabrerizo, F. J., Alonso, S., and Herrera-Viedma, E. 2009. "A consensus model for group decision making problems with unbalanced fuzzy linguistic information," *International Journal of Information Technology & Decision Making* (8:01), pp. 109–131.

Cabrerizo, F. J., Moreno, J. M., Pérez, I. J., and Herrera-Viedma, E. 2010. "Analyzing consensus approaches in fuzzy group decision making: advantages and drawbacks," *Soft Computing* (14:5), pp. 451–463.

Campello, R. J. 2007. "A fuzzy extension of the Rand index and other related indexes for clustering and classification assessment," *Pattern Recognition Letters* (28:7), pp. 833–841.

Carlile, P. R. 2002. "A pragmatic view of knowledge and boundaries: Boundary objects in new product development," *Organization science* (13:4), pp. 442–455.

Carlile, P. R. 2004. "Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries," *Organization science* (15:5), pp. 555–568.

Carlile, P. R., and Rebentisch, E. S. 2003. "Into the black box: The knowledge transformation cycle," *Management science* (49:9), pp. 1180–1195.

Carlsson, C., Ehrenberg, D., Eklund, P., Fedrizzi, M., Gustafsson, P., Lindholm, P., Merkuryeva, G., Riissanen, T., and Ventre, A. G. 1992. "Consensus in distributed soft environments," *European Journal of Operational Research* (61:1), pp. 165–185.

Chen, C.-T. 2000. "Extensions of the TOPSIS for group decision-making under fuzzy environment," *Fuzzy sets and systems* (114:1), pp. 1–9.

Chesbrough, H. 2006. "Open Innovation: A New Paradigm for Understanding Industrial Innovation," in *Open innovation: Researching a new paradigm*, H. Chesbrough, W. Vanhaverbeke, and J. West (eds.), Oxford university press, pp. 1–25.

Chesbrough, H. W., and Appleyard, M. M. 2007. "Open innovation and strategy," *California management review* (50:1), pp. 57–76.

Chiclana, F., Tapia García, J., del Moral, M. J., and Herrera-Viedma, E. 2013. "A statistical comparative study of different similarity measures of consensus in group decision making," *Information Sciences* (221), pp. 110–123.

Claxton, G. 2000. *Hare brain, tortoise mind: How intelligence increases when you think less*, {Harper Perennial}.

Cohen, J. 1968. "Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit." *Psychological bulletin* (70:4), p. 213.

Cohen, W. M., and Levinthal, D. A. 1990. "Absorptive capacity: A new perspective on learning and innovation," *Administrative science quarterly* pp. 128–152.

Cooper, R. G., and De Brentani, U. 1984. "Criteria for screening new industrial products," *Industrial Marketing Management* (13:3), pp. 149–156.

Dahlander, L., and Gann, D. M. 2010. "How open is innovation?" *Research policy* (39:6), pp. 699–709.

Dahlin, K. B., Weingart, L. R., and Hinds, P. J. 2005. "Team diversity and information use," *Academy of Management Journal* (48:6), pp. 1107–1123.

De Bono, E. 1985. "Six Thinking Hats: An Essential Approach to Business Management from the Creator of Lateral Thinking," *Little, Brown, & Company* .

Diehl, M., and Stroebe, W. 1987. "Productivity loss in brainstorming groups: Toward the solution of a riddle." *Journal of personality and social psychology* (53:3), p. 497.

Dou, W., Ren, Y., Wu, Q., Ruan, S., Chen, Y., Bloyet, D., and Constans, J.-M. 2007. "Fuzzy kappa for the agreement measure of fuzzy classifications," *Neurocomputing* (70:4), pp. 726–734.

Dubois, D., and Prade, H. 2004. "On the use of aggregation operations in information fusion processes," *Fuzzy Sets and Systems* (142:1), pp. 143–161.

Easley, D., and Kleinberg, J. 2010. *Networks, crowds, and markets: Reasoning about a highly connected world*, Cambridge University Press.

Ebner, W., Leimeister, J. M., and Krcmar, H. 2009. "Community engineering for innovations: the ideas competition as a method to nurture a virtual community for innovations," *R&d Management* (39:4), pp. 342–356.

Edmondson, A. 1999. "Psychological safety and learning behavior in work teams," *Administrative science quarterly* (44:2), pp. 350–383.

Edwards, W. 1977. "How to use multiattribute utility measurement for social decisionmaking," *Systems, Man and Cybernetics, IEEE Transactions on* (7:5), pp. 326–340.

Estellés-Arolas, E., and González-Ladrón-de Guevara, F. 2012. "Towards an integrated crowdsourcing definition," *Journal of Information Science* (38:2), pp. 189–200.

Evans, M. S. 2014. "A computational approach to qualitative analysis in large textual datasets," *PloS one* (9:2), p. e87,908.

Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. 2011. "Knowledge collaboration in online communities," *Organization science* (22:5), pp. 1224–1239.

Felin, T., and Zenger, T. R. 2011. "Information aggregation, matching and radical market–hierarchy hybrids: Implications for the theory of the firm," *Strategic Organization* (1), p. 11.

Felin, T., and Zenger, T. R. 2014. "Closed or open innovation? Problem solving and the governance choice," *Research Policy* (43:5), pp. 914–925.

Feller, J., Finnegan, P., Hayes, J., and O'Reilly, P. 2012. "'Orchestrating'sustainable crowdsourcing: A characterisation of solver brokerages," *The Journal of Strategic Information Systems* (21:3), pp. 216–232.

Ferguson, D. 2010. "Global Pulse 2010: Insights and ideas from around the world," Tech. rep., United States Agency for International Development.

Fowlkes, E. B., and Mallows, C. L. 1983. "A method for comparing two hierarchical clusterings," *Journal of the American statistical association* (78:383), pp. 553–569.

Franke, N., and Shah, S. 2003. "How communities support innovative activities: an exploration of assistance and sharing among end-users," *Research policy* (32:1), pp. 157–178.

Frey, K., Lüthje, C., and Haag, S. 2011a. "Whom should firms attract to open innovation platforms? The role of knowledge diversity and motivation," *Long Range Planning* (44:5), pp. 397–420.

Frey, K., Lüthje, C., and Haag, S. 2011b. "Whom should firms attract to open innovation platforms? The role of knowledge diversity and motivation," *Long Range Planning* (44:5), pp. 397–420.

Füller, J., Hutter, K., and Fries, M. 2012a. "Crowdsourcing for Goodness Sake: Impact of Incentive Preference on Contribution Behavior for Social Innovation," *Adv Int Market* (23), pp. 137–159.

Füller, J., Hutter, K., Hautz, J., and Matzler, K. 2014. "User Roles and Contributions in Innovation-Contest Communities," *Journal of Management Information Systems* (31:1), pp. 273–308.

Füller, J., Matzler, K., Hutter, K., and Hautz, J. 2012b. "Consumers' creative talent: Which characteristics qualify consumers for open innovation projects? An exploration of asymmetrical effects," *Creativity and Innovation Management* (21:3), pp. 247–262.

Füller, J., Möslein, K., Hutter, K., and Haller, J. 2010. "Evaluation Games-How to Make the Crowd your Jury." in *GI Jahrestagung (1)*, .

Gamer, M., Lemon, J., Fellows, I., and Singh, P. 2012. "irr: Various coefficients of interrater reliability and agreement," *R package version 084* .

Girotra, K., Terwiesch, C., and Ulrich, K. T. 2010. "Idea generation and the quality of the best idea," *Management science* (56:4), pp. 591–605.

Godart, F., Maddux, W., Shipilov, A., and Galinsky, A. 2014. "Fashion with a Foreign Flair: Professional Experiences Abroad Facilitate the Creative Innovations of Organizations," *Academy of Management Journal* pp. amj–2012.

Goddard, C. 2011. *Semantic analysis: A practical introduction*, Oxford University Press.

Gray, P. H. 2000. "The effects of knowledge management systems on emergent teams: towards a research model," *The journal of strategic information systems* (9:2), pp. 175–191.

Guggenmoos-Holzmann, I. 1996. "The meaning of kappa: probabilistic concepts of reliability and validity revisited," *Journal of clinical epidemiology* (49:7), pp. 775–782.

Gwet, K. L. 2014. *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*, Advanced Analytics, LLC.

Hahn, J., Moon, J. Y., and Zhang, C. 2008. "Emergence of new project teams from open source software developer networks: Impact of prior collaboration ties," *Information Systems Research* (19:3), pp. 369–391.

Halkidi, M., Batistakis, Y., and Vazirgiannis, M. 2001. "On clustering validation techniques," *Journal of Intelligent Information Systems* (17:2-3), pp. 107–145.

Hargadon, A. 2003. *How breakthroughs happen: The surprising truth about how companies innovate*, Harvard Business Press.

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

Harrison, D. A., Price, K. H., and Bell, M. P. 1998. "Beyond relational demography: Time and the effects of surface-and deep-level diversity on work group cohesion," *Academy of management journal* (41:1), pp. 96–107.

Herrera, F., Herrera-Viedma, E., and Verdegay, J. L. 1997. "Linguistic measures based on fuzzy coincidence for reaching consensus in group decision making," *International Journal of Approximate Reasoning* (16:3), pp. 309–334.

Herrera-Viedma, E., Alonso, S., Chiclana, F., and Herrera, F. 2007. "A consensus model for group decision making with incomplete fuzzy preference relations," *Fuzzy Systems, IEEE Transactions on* (15:5), pp. 863–877.

Herrera-Viedma, E., Cabrerizo, F. J., Kacprzyk, J., and Pedrycz, W. 2014. "A review of soft consensus models in a fuzzy environment," *Information Fusion* (17), pp. 4–13.

Herrera-Viedma, E., Herrera, F., and Chiclana, F. 2002. "A consensus model for multiperson decision making with different preference structures," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* (32:3), pp. 394–402.

Herrera-Viedma, E., Martinez, L., Mata, F., and Chiclana, F. 2005. "A consensus support system model for group decision-making problems with multigranular linguistic preference relations," *Fuzzy Systems, IEEE Transactions on* (13:5), pp. 644–658.

Homan, A. C., Hollenbeck, J. R., Humphrey, S. E., Van Knippenberg, D., Ilgen, D. R., and Van Kleef, G. A. 2008. "Facing differences with an open mind: Openness to experience, salience of intragroup differences, and performance of diverse work groups," *Academy of Management Journal* (51:6), pp. 1204–1222.

Howe, J. 2006. "The rise of crowdsourcing," *Wired magazine* (14:6), pp. 1–4.

Hox, J. J., Moerbeek, M., and van de Schoot, R. 2010. *Multilevel analysis: Techniques and applications*, Routledge.

Hubert, L., and Arabie, P. 1985. "Comparing partitions," *Journal of classification* (2:1), pp. 193–218.

Hutter, K., Hautz, J., Füller, J., Mueller, J., and Matzler, K. 2011. "Communitition: The tension between competition and collaboration in community-based design contests," *Creativity and Innovation Management* (20:1), pp. 3–21.

Jackson, S. E., May, K. E., Whitney, K., et al. 1995. "Understanding the dynamics of diversity in decision-making teams," *Team effectiveness and decision making in organizations* (204), p. 261.

Jain, A. K., Dubes, R. C., et al. 1988. *Algorithms for clustering data*, vol. 6, Prentice hall Englewood Cliffs.

Jeppesen, L. B. 2005. "User toolkits for innovation: Consumers support each other," *Journal of Product Innovation Management* (22:4), pp. 347–362.

Jeppesen, L. B., and Lakhani, K. R. 2010. "Marginality and problem-solving effectiveness in broadcast search," *Organization science* (21:5), pp. 1016–1033.

Jiang, D., Tang, C., and Zhang, A. 2004. "Cluster analysis for gene expression data: A survey," *Knowledge and Data Engineering, IEEE Transactions on* (16:11), pp. 1370–1386.

Joshi, A., and Roh, H. 2009. "The role of context in work team diversity research: A meta-analytic review," *Academy of Management Journal* (52:3), pp. 599–627.

Kacprzyk, J., and Fedrizzi, M. 1988. "A 'soft' measure of consensus in the setting of partial (fuzzy) preferences," *European Journal of Operational Research* (34:3), pp. 316–325.

Kane, G. C., Johnson, J., and Majchrzak, A. 2014. "Emergent life cycle: The tension between knowledge change and knowledge retention in open online coproduction communities," *Management Science* (60:12), pp. 3026–3048.

Kaner, S. 2014. *Facilitator's guide to participatory decision-making*, John Wiley & Sons.

Kavadias, S., and Sommer, S. C. 2009. "The effects of problem structure and team diversity on brainstorming effectiveness," *Management Science* (55:12), pp. 1899–1913.

Kim, A. J. 2000. *Community building on the web: Secret strategies for successful online communities*, Addison-Wesley Longman Publishing Co., Inc.

King, A. A., and Lakhani, K. R. 2011. *The contingent effect of absorptive capacity: An open innovation analysis*, Harvard Business School.

Kleining, G., and Witt, H. 2001. "Discovery as basic methodology of qualitative and quantitative research," in *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, vol. 2, , vol. 2.

Kukar, M. 2003. "Transductive reliability estimation for medical diagnosis," *Artificial Intelligence in Medicine* (29:1-2), pp. 81–106.

Kuppuswamy, V., and Bayus, B. L. 2015. "Crowdfunding creative ideas: The dynamics of project backers in Kickstarter," *UNC Kenan-Flagler Research Paper* (1:2013-15).

Lakhani, K., Garvin, D. A., and Lonstein, E. 2010. "Topcoder (a): Developing software through crowdsourcing," *Harvard Business School General Management Unit Case* (1:610-032).

Lakhani, K. R., Jeppesen, L. B., Lohse, P. A., and Panetta, J. A. 2007. *The value of openess in scientific problem solving*, Division of Research, Harvard Business School.

Lamastra, C. R. 2009. "Software innovativeness. A comparison between proprietary and Free/Open Source solutions offered by Italian SMEs," *R&D Management* (39:2), pp. 153–169.

Landis, J. R., and Koch, G. G. 1977. "The measurement of observer agreement for categorical data," *biometrics* pp. 159–174.

Laursen, K., and Salter, A. 2006. "Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms," *Strategic management journal* (27:2), pp. 131–150.

Leimeister, J. M., Huber, M., Bretschneider, U., and Krcmar, H. 2009. "Leveraging crowdsourcing: activation-supporting components for IT-based ideas competition," *Journal of management information systems* (26:1), pp. 197–224.

Leiponen, A., and Helfat, C. E. 2010. "Innovation objectives, knowledge sources, and the benefits of breadth," *Strategic Management Journal* (31:2), pp. 224–236.

Leonard, D., and Rayport, J. F. 1997. "Spark innovation through empathic design," *Harvard business review* (75), pp. 102–115.

Leonard-Barton, D. 1998. *Wellsprings of knowledge: Building and sustaining the sources of innovation*, Harvard Business Press.

Leuf, B., and Cunningham, W. 2001. *The Wiki way: quick collaboration on the Web*, vol. 9, Addison-Wesley Boston.

Lim, S. L., Quercia, D., and Finkelstein, A. 2010. "StakeSource: harnessing the power of crowdsourcing and social networks in stakeholder analysis," in *Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering-Volume 2*, , ACM.

Linden, S. 2014. "How Crowdsourcing Can Reduce the Reliability of Social Media Analytics," Tech. rep., Gartner.

Linstone, H. A., Turoff, M., et al. 1975. *The Delphi method: Techniques and applications*, vol. 29, Addison-Wesley Reading, MA.

Locke, K. D., and Horowitz, L. M. 1990. "Satisfaction in interpersonal interactions as a function of similarity in level of dysphoria." *Journal of personality and social psychology* (58:5), p. 823.

Lukyanenko, R., Parsons, J., and Wiersma, Y. F. 2014. "The IQ of the crowd: understanding and improving information quality in structured user-generated content," *Information Systems Research* (25:4), pp. 669–689.

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

Majchrzak, A., and Malhotra, A. 2013. "Towards an information systems perspective and research agenda on crowdsourcing for innovation," *The Journal of Strategic Information Systems* (22:4), pp. 257–268.

Majchrzak, A., More, P. H., and Faraj, S. 2012a. "Transcending knowledge differences in cross-functional teams," *Organization Science* (23:4), pp. 951–970.

Majchrzak, A., Wagner, C., and Yates, D. 2012b. "The impact of shaping on knowledge reuse for organizational improvement with wikis," *MIS Quarterly, February* .

Majchrzak, A., Wagner, C., and Yates, D. 2013. "The Impact of Shaping on Knowledge Reuse for Organizational Improvement with Wikis." *Mis Quarterly* (37:2), pp. 455–469.

Malhotra, A., and Majchrzak, A. 2014. "Managing Crowds in Innovation Challenges," *California Management Review* (56:4).

Malhotra, A., Majchrzak, A., and Niemiec, R. M. 2016. "Using Public Crowds for Open Strategy Formulation: Mitigating the Risks of Knowledge Gaps," *Long Range Planning* .

Martínez, L., and Montero, J. 2007. "Challenges for improving consensus reaching process in collective decisions," *New Mathematics and Natural Computation* (3:02), pp. 203–217.

Maznevski, M. L. 1994. "Understanding our differences: Performance in decision-making groups with diverse members," *Human relations* (47:5), pp. 531–552.

McGrath, R. G., Tsai, M.-H., Venkataraman, S., and MacMillan, I. C. 1996. "Innovation, competitive advantage and rent: a model and test," *Management Science* (42:3), pp. 389–403.

McIntyre, R. M. M. M. M. C., A. 2013. "Predicts 2014: Consumer Analytics and Personalized User Experiences Transform Competitive Advantage." Tech. rep., Statistical techniques for the study of language behaviour.

Nahapiet, J., and Ghoshal, S. 1998. "Social capital, intellectual capital, and the organizational advantage," *Academy of management review* (23:2), pp. 242–266.

Neyer, A.-K., Bullinger, A. C., and Moeslein, K. M. 2009. "Integrating inside and outside innovators: a sociotechnical systems perspective," *R&d Management* (39:4), pp. 410–419.

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

OpenIDEO 2015-02-03. "OpenIDEO - Where people design better, together," ttps://openideo.com/.

Osei-Bryson, K.-M. 2010. "Towards supporting expert evaluation of clustering results using a data mining process model," *Information Sciences* (180), pp. 414–431.

Page, S. E. 2008. *The difference: How the power of diversity creates better groups, firms, schools, and societies*, Princeton University Press.

Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L., and Roberts, R. D. 2002. "The role of individual differences in the accuracy of confidence judgments," *The Journal of General Psychology* (129:3), pp. 257–299.

Palomares, I., Estrella, F. J., Martínez, L., and Herrera, F. 2014. "Consensus under a fuzzy context: Taxonomy, analysis framework AFRYCA and experimental case of study," *Information Fusion* (20), pp. 252–271.

Paulus, P. B., and Nijstad, B. A. 2003. *Group creativity: Innovation through collaboration.*, Oxford University Press.

Pierce, J. L., and Delbecq, A. L. 1977. "Organization structure, individual attitudes and innovation," *Academy of management review* (2:1), pp. 27–37.

Pieterse, A. N., Van Knippenberg, D., and van Ginkel, W. P. 2011. "Diversity in goal orientation, team reflexivity, and team performance," *Organizational Behavior and Human Decision Processes* (114:2), pp. 153–164.

Piller, F. T., and Walcher, D. 2006. "Toolkits for idea competitions: a novel method to integrate users in new product development," *r&D management* (36:3), pp. 307–318.

Puranam, P., Gulati, R., and Bhattacharya, S. 2013. "How much to make and how much to buy? An analysis of optimal plural sourcing strategies," *Strategic Management Journal* (34:10), pp. 1145–1161.

Rand, W. M. 1971. "Objective criteria for the evaluation of clustering methods," *Journal of the American Statistical association* (66:336), pp. 846–850.

Red Cross, I. C. o. t. 2015-02-03. "How Does co-creation Change the Approach People are Taking to Innovation?" http://blogs.icrc.org/gphi2/2014/12/19/co-creation/.

Rietveld, T., and Van Hout, R. 1993. *Statistical techniques for the study of language and language behaviour*, Walter de Gruyter.

Rittel, H. W., and Webber, M. M. 1973. "Dilemmas in a general theory of planning," *Policy sciences* (4:2), pp. 155–169.

Ruspini, E. H. 1969. "A new approach to clustering," *Information and control* (15:1), pp. 22–32.

Saldaña, J. 2012. *The coding manual for qualitative researchers*, 14, Sage.

Santos, R., and Spann, M. 2011. "Collective entrepreneurship at Qualcomm: combining collective and entrepreneurial practices to turn employee ideas into action," *R&D Management* (41:5), pp. 443–456.

Sawhney, M., Verona, G., and Prandelli, E. 2005. "Collaborating to create: The Internet as a platform for customer engagement in product innovation," *Journal of interactive marketing* (19:4), pp. 4–17.

Schulze, A., and Hoegl, M. 2008. "Organizational knowledge creation and the generation of new product ideas: A behavioral approach," *Research policy* (37:10), pp. 1742–1750.

Schumpeter, J. A. 2013. *Capitalism, socialism and democracy*, Routledge.

Simonton, D. K. 1999. *Origins of genius: Darwinian perspectives on creativity.*, Oxford University Press.

Surowiecki, J. 2005. *The wisdom of crowds*, Anchor.

Tausczik, Y. R., Kittur, A., and Kraut, R. E. 2014. "Collaborative problem solving: a study of mathoverflow," in *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, , ACM.

Taylor, A., and Greve, H. R. 2006. "Superman or the fantastic four? Knowledge combination and experience in innovative teams," *Academy of Management Journal* (49:4), pp. 723–740.

Teachman, J. D. 1980. "Analysis of population diversity measures of qualitative variation," *Sociological Methods & Research* (8:3), pp. 341–362.

Terwiesch, C., and Ulrich, K. T. 2009. *Innovation tournaments: Creating and selecting exceptional opportunities*, Harvard Business Press.

Terwiesch, C., and Xu, Y. 2008. "Innovation contests, open innovation, and multiagent problem solving," *Management science* (54:9), pp. 1529–1543.

Tierney, M. L., and Drury, J. 2013. "Continuously Improving Innovation Management through Enterprise Social Media," *Journal of Social Media for Organizations* (1:1), pp. 1–16.

Tiwana, A. 2008. "Do bridging ties complement strong ties? An empirical examination of alliance ambidexterity," *Strategic Management Journal* (29:3), pp. 251–272.

Tiwana, A., and Mclean, E. R. 2005. "Expertise integration and creativity in information systems development," *Journal of Management Information Systems* (22:1), pp. 13–43.

Torrance, G. W., Boyle, M. H., and Horwood, S. P. 1982. "Application of multi-attribute utility theory to measure social preferences for health states," *Operations research* (30:6), pp. 1043–1069.

Toubia, O. 2006. "Idea generation, creativity, and incentives," *Marketing Science* (25:5), pp. 411–425.

Tsoukas, H. 2009. "A dialogical approach to the creation of new knowledge in organizations," *Organization Science* (20:6), pp. 941–957.

Tushman, M. L. 1997. "Winning through innovation," *Strategy & Leadership* (25:4), pp. 14–19.

Ulrich, K. T. 2011. *Design: Creation of artifacts in society*, Karl T. Ulrich.

Van Der Vegt, G. S., Bunderson, J. S., and Oosterhof, A. 2006. "Expertness diversity and interpersonal helping in teams: Why those who need the most help end up getting the least," *Academy of Management Journal* (49:5), pp. 877–893.

Van Knippenberg, D., and Schippers, M. C. 2007. "Work group diversity," *Annu Rev Psychol* (58), pp. 515–541.

van Osch, W., and Avital, M. 2010. "Generative collectives," in *ICIS*, .

Vanhaverbeke, W., and Cloodt, M. 2015. "Theories of the firm and open innovation," in *New Frontiers in Open Innovation*, , Cambridge, UK: Oxford University Press, p. 256.

Villarroel, J. A., Taylor, J. E., and Tucci, C. L. 2013. "Innovation and learning performance implications of free revealing and knowledge brokering in competing communities: insights from the Netflix Prize challenge," *Computational and Mathematical Organization Theory* (19:1), pp. 42–77.

Von Hippel, E. 1989. *Cooperation between rivals: informal know-how trading*, Springer.

Von Hippel, E. 2005. "Democratizing innovation: The evolving phenomenon of user innovation," *Journal für Betriebswirtschaft* (55:1), pp. 63–78.

Von Hippel, E. 2007. *The sources of innovation*, Springer.

von Hippel, E., and von Krogh, G. 2003. "Open source software and the "private-collective" innovation model: Issues for organization science," *Organization science* (14:2), pp. 209–223.

Wang, C., Rodan, S., Fruin, M., and Xu, X. 2013. "Knowledge networks, collaboration networks, and exploratory innovation," *Academy of Management Journal* pp. amj–2011.

West, J. 2009. "Policy Challenges of Open, Cumulative, and User Innovation," *Wash UJL & Pol'y* (30), p. 17.

West, J., and Bogers, M. 2014. "Leveraging external sources of innovation: a review of research on open innovation," *Journal of Product Innovation Management* (31:4), pp. 814–831.

West, J., and Lakhani, K. R. 2008. "Getting clear about communities in open innovation," *Industry and Innovation* (15:2), pp. 223–231.

West, M. A. 2012. *Effective teamwork: Practical lessons from organizational research*, John Wiley & Sons.

Zadeh, L. A. 1965. "Fuzzy sets," *Information and control* (8:3), pp. 338–353.

Zaraté, P., Belaud, J. P., and Camilleri, G. 2008. *Collaborative decision making: Perspectives and challenges*, vol. 176, IOS Press.