Exploring Social Influence in Crowdsourcing

Candidate:

Rui Gonçalves Garrido de Lima Martins

ruigmartins@gmail.com

Advisor:

Prof. Juan Andrei Villarroel Fernández

andreiv@ucp.pt

September – 2012

ABSTRACT

This study addresses the phenomenon of social influence in crowdsourcing platforms and, in particular, it sought to explore whether depending on personal, contextual or cultural, characteristics, individuals which participate in such endeavors are more or less prone to become influenced by a majority when performing an objective task.

In crowdsourcing platforms individuals participate through different decision making mechanisms such as voting, averaging and consensus. The exposure to informational cues regarding the aggregate output of a crowd is hypothesized to constitute a pressure upon the individuals' behavior.

Our findings provide evidence about the existence of and proneness to social influence in crowdsourcing platforms. These research also shows that skill level and *homophilous* ties among participants are a stronghold against the pressures of the majorities. These results empirically demonstrate that in the presence of an informational cues a large amount of individuals accept majorities' outcomes as an evidence of the reality.

Through the current findings, we provide insight on the phenomenon of social influence in crowdsourcing platforms and, by unveiling these results, give a contribution towards managing such platforms.

Table of Contents

ACKNOWLEDGMENTS

I would like to thank my thesis advisor, Prof. Andrei Villarroel, for his invaluable mentorship, infinite patient, and critical guidance. His important support made this learning experience fruitful.

To my colleagues in the Online Distributed Organization dissertation seminar, for their valuable remarks in the discussion sessions and for their contributions to this study.

To Noémia, I am deeply grateful for her unconditional help and for having been an endless source of motivation.

To my parents and sister, for being always present with a message of encouragement.

To all of those who participated in this study. Your contribution was widely appreciated.

I – INTRODUCTION

The main goals of this investigation is to understand if social influence is present in crowdsourcing and hence to deepen the knowledge about working with these platforms; namely, exploring how individuals in an online crowd behave when they are subject to pressures from a majority. Social influence is in this research a concept of paramount importance that we define as a convergence or alignment of thoughts and actions to match the responses of others.

In both non-intra and intra-organizational crowdsourcing platforms, individuals have similar group decision mechanisms (e.g. vote, averaging, consensus), which they commonly use to contribute to the online communities they belong. Of direct relevance to the present study, is that in online crowdsourcing platforms, individuals are typically exposed to informative cues, such as the number of votes a particular solution has received.

Our premise is that such crowdsourcing platforms mediate social processes and the existence of informative cues about the general opinion of the community could constrain individual contributions because they can materialize an influence in a form of dominance (Asch, 1955). In such conditions the outcomes of the participation are undermined, individuals become less independent and thus the social dynamic could be detrimental to the goal pursued.

To attain our objective, we asked individuals in the crowd to perform a set of problem-solving objective tasks and we introduced some informative cues to test whether they were permeable to the pressures of others. Moreover, inquiries were made about the circumstances in which this social phenomenon has more impact on the performance of individuals. We controlled for how social influence may also depend on characteristics of the crowd employed, as well as cultural and demographic factors known to influence social behavior.

Through the current study we shed light over this social phenomenon in crowdsourcing platforms. As crowdsourcing is emerging as a new organizational model with successful results in the managerial world, this contribution is useful to

unveil this singularity and to untangle the circumstances in which social influence unsettle the wisdom of the crowds.

Although studies about social influence have initially emphasized behavioral observations of individuals in face to face situations (Asch, 1955; Deutsch and Gerard, 1955); mathematical models which tried to anticipate the level influence a majority can create (Latané, 1981; Tanford and Penrod, 1984); or, more recently, the individual susceptibility to be influenced by a group in a context of computer mediated communications (Postmes et al., 1998); few, if not nonexistent, were the efforts to understand this social dynamic the scope of crowdsourcing platforms.

To achieve this aim, firstly we contextualize this study within the scope of literature that exists about crowdsourcing (section 2.1). Later, we review the most relevant streams of literature that address the social influence phenomenon (sections 2.2, 2.3 and 2.4). A particular emphasis was given to the most relevant works on social psychology and other fields concerned with social networks as channels to spread of behavior. We end this chapter by raising the hypotheses that our study was set to test (section 2.5).

Secondly, the methodological context for this study is presented. At this stage the data analysis is presented, as well as the variables that were used throughout the last part of our study (sections 3.5 and 3.6).

Lastly, we devote our attention to discussing the results of the previous analysis (sections 4.1, 4.2 and 4.3), present the main findings, establish the limitations of the study (section 5.2), and suggest some directions for future research (section 5.3).

II – LITERATURE REVIEW

In order to display our hypotheses testing and proceed with the analysis, in the following sections are displayed the relevant streams of literature. We start by contextualizing this investigation within the crowdsourcing boundaries and thereafter the emphasis is placed upon the literature that address the phenomenon of social influence. At this point, we considered two distinctive, but complementary,

flows of literature: the analysis starts firstly with the relevant developments on social psychology and after scrutinizes what was written about social networks as channels to spread similar behaviors.

2.1 Crowdsourcing and organization

The crowdsourcing concept was informally coined by Jeff Howe (2006), and etymologically stems at the contraction of the word "Crowd" and "Outsourcing". In accordance to Howe's definition, crowdsourcing is an opportunity companies have to outsource directly to a crowd, rather than a company. The concept emphasizes the notion of the open call, which is the possibility to broadcast problems or challenges to wide range of potential solvers or contributors.

Crowdsourcing as an organizational model is a relatively new domain of academic inquiry (Villarroel and Tucci, 2007; Lakhani and Panetta, 2007; Villarroel, 2008) which is found at the intersection of different fields of research (Villarroel 2008: Chapter 1; Schenk and Guittard, 2009). The managerial relevance of the concept is demonstrated by successful organizations such as Facebook, Google, Wikipedia, which, in 2012, were among the top five with the most web traffic worldwide[1], indicative of the value of working with crowds (Villarroel, 2008, 2011). At the same time crowdsourcing is starting to play a role on public policy. For this purpose, President Obama's signing of COMPETES act[2] depicts the potential crowdsourcing has to bridge new ideas and generate the breakthroughs of tomorrow.

Crowdsourcing is partly a consequence of a new era in which new computer capabilities and electronic networks grew exponentially. The increasing division of labor in societies and the explosion of knowledge in diverse scientific and

---

1 In accordance to Alexa – The Web Information Company (http://www.alexa.com/topsites). Last accessed on July 19, 2012.

[2] In January 2011, President Obama signed into a law that granted federal agencies authority to conduct challenges in the form of open competitions to spur innovation and solve tough problems through novel approaches.

technological fields imposed one critical challenge in the managerial world: harnessing individuals with expertise and thus overcome the fundamental problem of having relevant knowledge residing outside the boundaries of the firm.

Villarroel (2008: Chapter 4) defined the contours of crowdsourcing endeavors by separating them into two major groups: self-organized (e.g. Wikipedia) and firm-sponsored (e.g. Netflix Prize). Additionally, it was suggested that these activities range from collaborative to competitive and their implementation depends on the level of complexity of task (Villarroel 2008: Chapter 4).

This new paradigm comes in the wake of what Hayek (1945) suggested more than sixty years ago as the knowledge unevenly distributed in societies. His investigations about the nature of the problem of a rational economic order led him to believe that it was «determined precisely by the fact that the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form, but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess» (Hayek; p: 519).

Although this contribution allowed to foresee the implications of having resources spread throughout societies and to extrapolate its implications to the managerial world, crowdsourcing also found its roots in two streams of literature: Open Source Software and User Innovation. While the former emphasized endeavors (e.g. Linux) in which there was an intrinsic benefit of freely revealing innovations (von Hippel & von Krogh, 2003); the latter was focus on the fact that users, rather than companies, are the sources of innovation because they experience specific needs (von Hippel, 2001). In accordance, at the intersection of the aforementioned fields of research, crowdsourcing emerges as a novel paradigm which seek to mobilize the living competence and expertise of the crowd (Schenk and Guittard, 2009).

Research upon the nature of crowdsourcing tasks were separated between complex, creative and routine tasks (Schenk & Guittard, 2009). The first of the aforementioned set of tasks was described to be useful when companies lack the

in-house satisfactory solutions (Lakhani & Panetta, 2007). In such situations, they may turn to a crowd and ask for their problem solving skills. Differently, individuals can also work upon creative tasks and thus the broadcast search is aimed to obtain novel contributions. Activities ranging from image analysis to nature language processing, what is generally named artificial intelligence problems (Villarroel & Tucci, 2010) or routine tasks (Schenk & Guittard, 2009), are another type of jobs usually found in crowdsourcing platforms. Outsourcing such tasks is the consequence of inabilities computers have to process tasks that for humans are very easy to solve. Although they are, in opposition to complex and creative tasks, poor from a cognitive point of view, the advantage of having a crowd performing such tasks rely on the low cost and the large scale (Schenk & Guittard, 2009).

Some of the advantages of this paradigm were already pointed out. Villarroel (2011) asserted that crowdsourcing is a new form of organization which help to outperform traditional organizations due to the following advantages:

- Market reach – Due to an unprecedented global reach and an inherent scale factor advantage;

- Quality - The quality of the inputs created by the crowds were considered to be of remarkable quality. Nature magazine (2005) published an article about the science entries on Wikipedia being "head to head" with encyclopedia Britannica;

- Service responsiveness - Due to computing and communication capabilities increasingly responsive, crowds are able to work in an extremely effective way and correct promptly any errors that may exist;

- Cost effectiveness - Companies are allowed to access knowledge at low (e.g. Amazon Mechanical Turk), or at relatively low (e.g. Innocentive), cost. Facebook Translations, a 2007 crowdsourcing initiative aimed to translate the actual corpus powering Facebook website, in less than a year engaged more than 300,000 people and translated the social network in more than 70 different languages at no cost (Villarroel, 2011).

This emerging model is however subject to some contrarieties. The disadvantages usually assigned concern the less privacy and less appropriability of intellectual property (Villarroel, 2011; Villarroel and Reis 2010). Both concerns are mainly due to the fact that crowdsourcing requires some level of transparency and hence companies are forced to share and to become more open. In particular many companies may not find any advantage with weakening appropriability regimes when they have poor downstream asset positions (Pisano, 2006).

In such cases an alternative proposition to overcome management information flows is made through internal knowledge markets (Benbya & van Alstyne, 2011). These environments are characterized by being protected from external agents to the firm and operationalized by trading knowledge via price mechanisms. One of the main features of these platforms rely on bridging employees who were previously geographically and hierarchically distant to the firm's innovation process (Villarroel & Reis, 2010, 2012). As such, Villarroel and Reis saw this linking process as an intra-organizational crowdsourcing mechanism as it «extend problem-solving to a large and diverse pool of self-selected contributors, referred to as the crowd, beyond the formal internal boundaries of a multi-business firm» (2012, p. 6). These knowledge markets «cause resources to speak up and self-identify» (Benbya & van Alstyne, p. 66) and assume different forms, such as: forecasting via prediction markets; idea creation and evaluation through idea markets; problem-solving through innovation markets; and peer assistance through knowledge markets.

2.2 Social influence through the lenses of Social Psychology

The impact that a group of people has on one person's beliefs and actions has been profusely studied for decades by researchers in different domains of knowledge (Raafat et al., 2009). One of those domains was social psychology which developed and adopted different perspectives to understand this phenomenon of social change.

A staunch critic of human suggestibility, Asch (1955) tried to test if in previous studies in which participants were invited to form opinions and preferences on diverse matters, the changes thereof in the course of time were, in fact, experimental victories. He was particularly interested in understanding if in an objective, instead an ambiguous, task, participants would remain independent given the groups' inaccurate judgments. Contrary to his primary expectations, his findings contributed to clarify the effectiveness of majorities' pressures upon individuals through line-judgment conformity experiments.

Along with studies about obedience (Milgram, 1973) and compliance without an explicit social pressure (Freedman & Fraser, 1966), they established the foundations of research on social influence and, in particular, stimulated the research in two core dimensions: compliance and conformity (Cialdini & Goldstein, 2004).

Following the groundbreaking evidences that, in different circumstances, when ordinary people face pressures upon their behavior they change it in way they would not commonly do, researchers attempt to scrutinize the factors that modify the level of social influence. Deutsch and Gerard (1955) made a first effort to separate the different motivations beyond social conformity. After considering several studies, they distinguished two types of motivation to conform: normative and informative social influence. They describe the former as an influence targeted to an individual to conform accordingly to the positive expectation of a group; whereas the latter is the motivation to accept information from another source due to the objective of producing accurate interpretations of the reality. Although they are commonly difficult to separate (Cialdini & Goldstein, 2004), some researchers have suggested that while informational mechanisms were assessed to have more impact in an ambiguous situation when it is difficult to make a clear judgment; normative social influence is suggested to be more involved when a condition is obviously wrong and conformity exist to avoid unpleasant feeling of a majority (Campbell & Fairey, 1989).

In addition to these early conceptualizations, researchers later scrutinized other aspects that could alter the effectiveness of social influence. The size of the majority was regarded to be among the critical factors to induce behavior replication (Asch, 1955, Morris & Miller, 1975). In Asch's face-to-face studies, evidences show that when an individual is confronted with an opposition of solely one individual, he or she is weakly persuaded. However, when the opposition accounts for two or three individuals, pressures toward conformity reach its maximum. Similarly, Gerard, Wilhelmy and Conolley's (1968) saw the size of the majority as a determinant variable towards a higher or lesser conformity. In their investigations on 154 high school students, the number of confederates providing incorrect judgments ranged between 1 and 7 and there was a linear increase of influence as long as the majority size increased. Other writers claimed the importance of this variable through mathematical models (Cialdini & Trost, 1998; Latané, 1981; Tanford and Penrod, 1984). The Theory of Social Impact (Latané, 1981) suggested that the amount of influence is the result of a multiplicative function of three variables: strength of the source to a target; proximity in time and space; and the number of sources influencing the target. Similarly, the Social Influence Model (Latané et al., 1984) suggests the number of the sources of influence as the most relevant predictor of conformity, not disparaging the importance of the consistency of the influence and task type.

This latter factor was in fact an issue of early discussion and was empirically tested in Asch's seminal studies, due to his doubts regarding previous results with ambiguous tasks. In this vein, research on social conformity developed hypothesis about the nature of the task and its level of difficulty as a predictor of social influence. At this purpose, findings suggest that pressures towards conformity are more easily found in non-factual oriented tasks, whereas factual task displayed a less pronounced effectiveness (Blake, Helson, & Mouton, 1957). In contrast, as long the task is more difficult to judge, pressures towards conformity become higher. Therefore, when an individual is convinced about what is the right judgment

to make in a particular situation, he or she is more prone to resist to majorities' pressures (Coleman, Blake, & Mouton, 1958).

Of similar importance to increase the level of pressure towards a default outcome is the unanimity of the vote. It has been argued that as long as a minority starts to have more supporters on its side, individuals become more prone to examine the stimulus and the influence begins to wane (Asch, 1955; Cialdini & Trost, 1998). As such, the existence of dissenter, that is a person who is instructed to disagree with both the source and the target of influence, can diminish the weight of the influence (Asch, 1955). Additionally, it was suggested that, when a minority persists and is sure about its judgments, a target starts a validation processes which consists in the comparison of both the majority and minority sources of influence, which can result in different internal and public outcomes (Moscovici & Personnaz, 1979). In this sense, although publicly accepting the view of the majority, an individual can dissent internally with the aim of avoiding negative sanctions from the majority.

Although evidences suggest that deviant individuals undermine the effectiveness of pressures towards conformity, altering group cohesion was assessed to change the levels of social influence. In an attempt to increase the level of group commitment, Deutsch and Gerard (1955) added a reward to groups which perform better in a line judgment task. Their results supported the fact that normative social influence will have more effect upon individuals that are part of a group, when compared to those that does not have any sense of belonging. In addition, when comparisons were made between anonymous and face-to-face interactions, findings suggest a lessening of social influence when individuals don't know each other's.

In a different sense social psychologists regarded conformity in CMC, an acronym of computer mediated communications (Cialdini & Goldstein, 2004). The evidences were derived from new social dynamics found through the use of e-mail, voice mail, and desktop video conferencing. The central argument refers that the widespread introduction of computers in everyday life enable the creation of

computer systems that heavily connect people and organizations. In result, as these communications are mediated by computers, computer networks were viewed as inherently social (Wellman, 2001) and capable of altering social identities and social context towards relative anonymity, compared to more individuated forms of communication (Postmes et al., 1998). As such, one might expect that as long as the contributions of individuals through CMC are difficult to individuate, normative influence would have a weak impact. Although in CMC is generally postulated that individuals are free to choose their own agendas and interests, evidences has shown that in visually anonymous CMC it appears to increase the level of conformity to group norms (Postmes et al.,1998). This contradictory result is explained by what was named social identity model of deindividuation effects (SIDE) due to «the absence of individual cues enhances awareness of the group dimensions of identity and interaction, and reliance on related norms and standards» (Postmes et al., p. 703). Moreover, researchers stress that when group members are identifiable and their social identity is salient they were less prone to engage in an ingroup favoritism. Conversely, when they were deindividuated they started to strongly support their new ingroup. In sum, this stream of research argue that CMC have implications on social processes because it accentuates conformity in new social boundaries, similar to what is found in the real world. However, an exception exist when individuals do not identify to the norms of the group. In such situations they will not respect those boundaries and deindividuation is not going to result in social influence (Postmes et al., 1998).

2.3 Cultural Differences and Social Influence

Researchers were also focused on whether cultural factors affect the level of conformity experienced in different countries (Triandis et al. 1988; Bond & Smith, 1996).

Conformity to a group was sometimes viewed as to be dependent on Hofstede's (1980) generic division between collectivist and individualistic countries. This researcher was initially intrigued by how people in different countries perceived

and interpret the world. Among several dimensions, he looked upon individualism and how persons are active agents, who expect to take care of themselves and their immediate relatives and are not necessarily dependent on the organization they belong; in opposition to individuals which value a tightly-knit framework in society in which they expect others to take care of them in exchange of unconditional loyalty. As such, the position of a specific society within the scope of this dimension was dependent on its inclination towards "I" and "we" (individualistic and collectivistic society, respectively).

This relationship was due to the fact that «an essential attribute of collectivist cultures is that individuals may be induced to subordinate their personal goals to the goals of some collective, which is usually a stable ingroup (…), and much of the behavior of individuals may concern goals that are consistent with the goals of this ingroup» (Triandis et.al, p. 324). As a consequence, in collectivist cultures these ingroups tend to be stable even if they require costly demands from the individuals. In contrast, in individualistic countries individuals are faithful to more ingroups but they drop from them as long as they demand inconvenient behaviors and hence they form new in-groups. In result, «in a collectivistic culture, conformity does not have the same negative implications as in more individualistic countries» (Cialdini & Trost, p. 168) because individuals are more willing to achieve and maintain social harmony than to have individual success. At this purpose, meta-analysis of conformity studies using Asch-type line judgment task found that Hofstede's (1983) measures of Individualism – Collectivism significantly explain the levels of conformity (Bond & Smith, 1996).

2.4 Social Networks as Channels to Spread Behavior

Another stream of research looked at this phenomenon through a different approach. Instead of looking to transmission mechanisms with the aim of understanding how do individuals broadcast and replicate behavior, the emphasis was rather on individuals as units of well-defined modes of interactions which yield herding behavior (Raafat et al., 2009). In this sense, the scope of the investigations

emphasizes the social network, rather than the individual, as the source of behavioral diffusion across the population (Christakis & Fowler, 2008; Centola, 2010; Centola, 2011).

In this sphere two competing views prevail and found significant empirical support. The first stress that behavior diffuse across population due to what was called the strength of weak ties or "small world problems" (Travers & Milgram, 1969; Grannovetter, 1973). The starting point is that weak ties, as they link distinctive sectors in a network, act as "local bridges" which are extremely effective in spreading contagious. This characteristic is mainly due to the lack of redundancy that weak, in opposition to strong, ties display. As such, anything that is intended to be diffused (e.g. a rumor) can transverse a higher social distance and reach a larger number of people if the ties are not highly clustered (Grannovetter, 1973). In this regard, due to the existence of shortcuts infectious diseases were predicted to spread easily and quickly among networks that are less clustered (Watts & Strogatz, 1998).

In contrast, a different approach sustain that social behavior is subject to a complex contagion and hence individuals need to have a systematic reinforcement before convincing someone to adopt a behavior. In short, this alternative understanding refers that clustered networks provide more effectiveness on social reinforcement for adoption because they have more redundant ties. Centola (2010) confronted in a network two patterns of connectedness in which individuals were embedded: clustered-lattice network (high level redundant ties) and random network (low level of redundant ties). His aim was to understand in which of the settings the pattern of health behavior adoption was most noticed. He found that behavior and respective likelihood of adoption, as well as the level of commitment, were significantly higher in clustered networks. More striking, the author found that clustered networks are capable of expanding into a large-scale diffusion in a short period of time. At the same purpose, others researchers found that the spread of happiness was dependent on the degree of separation (one's friends' friends' friends) between sources and target of happiness, and those surrounded by many

happy people are expected to become happy in the future (Fowler & Christakis, 2008).

Although the level of proximity between individuals in social networks was found to be an explanation for spreading similar behaviors, they were also regarded as dependent on the sociodemographic variables or the degree of homophily (Centola, 2011). The latter is a principle that suggests that «a contact between similar people occurs at a high rate than among dissimilar people» (McPherson et al., p. 416). In accordance, variables such values, attitudes, age, ethnicity, education or occupation, when shared by individuals, were assumed to shape future patterns of behavior. Homophily was regarded to have an impact on the adoption of health behavior, as in homophilous conditions individuals were more prone to adopt a new type of behavior (Centola, 2010). Similarly, a study about the factors that induce the spread of obesity upon a network during 32 years reported that pairs of friends which share the same gender are more susceptible to become obese (Christakis & Fowler, 2007).

2.5 Research Hypothesis

The previous review of the literature unveiled some of the determinants factors that affect the level of social influence, as well as the spread of behavior in social networks. Accordingly, we expect that some of the factors will also intervene in crowdsourcing platforms.

In this sense, through the current investigation we try to understand if social influence exists and in what circumstances it impacts individuals on crowdsourcing platforms.

H1. Social influence will affect individual participants in crowdsourcing platforms.

A critical issue to this study is whether it is verifiable the existence of social influence in a crowdsourcing platform. Due to previous studies, our hint is that when participants are subject to external information about the outcome of a task they will

generally be prone to accept a cue as an evidence of the reality. At the same time, the existence of an influence will distract participants from giving a correct answers to an objective task and thus undermine their performance.

H1.a Social influence will be greater when individuals have ties among themselves.

Although our primary expectation refers that social influence will have a role to play in crowdsourcing platforms, we hypothesize that in those platforms where individuals have ties among themselves the effect will be more noticeable. As such, we depart from the assumption that when individuals share a common basis of identification (e.g. education, age, occupation, etc.), they will become more likely to behave similarly to participants which share the same traits. In this sense, if the source of influence share some similarities with the target, the pressures towards conformity will be most felt. In contrast, when individuals (source and target of influence) don't share any common basis, the pressures toward conformity will be less effective.

H1.b Social influence will be lesser when individuals are more skillful.

In accordance with the literature that suggest that task nature and difficulty has a significant impact on the susceptibility to conform (Blake, Helson, & Mouton, 1957), we hypothesize about the skills of the individuals being a stronghold to resist the pressures of the majorities in crowdsourcing platforms.

H2. Individuals for individualistic countries are less susceptible to accept informative cues in crowdsourcing platforms than participants from collectivist countries.

Cultural variables were also viewed as determinant factors to predict conformity to a group of people. As such, we asserted that, depending on the higher or lesser degree of individualistic/collectivist culture, individuals will be more or less prone to accept and validate external sources of information. In this sense, whereas

individuals from collectivist countries are more likely to accept external cues due to a cultural subordination towards the group norms, the opposite is predicted to happen with individuals from individualistic countries.

III – METHODOLOGY

I received a research design from Prof. Villarroel, describing a procedure to compare the performance of individual members of a crowd in two scenarios: (1) individuals in a strictly anonymous platform where no participant knows any other, (2) individuals in a known social network that all participants clearly relate to and where most of them know each other.

In order to approach the issues proposed we asked individuals to solve a set of problem-solving questions. To influence participants an informative cue was included in each question, henceforth called bias. The anonymous part of the test was conducted on Amazon Mechanical Turk (AMT), whereas the social part was performed with students from a top business school in Europe. AMT is a crowdsourcing platform used in research and it is considered to be valid for conducting behavioral, problem-solving and learning experiments, given that the results are in line with laboratory results (McDonnell et al., 2012). Earlier research by Prof. Villarroel had also been conducted on AMT – studying AMT itself – which informed and provided the basis for our present work (Villarroel, 2008: Chapter 4).

The study was conducted between April and July 2012, consisting of three steps described below.

3.1 Assessment of the Level of Difficulty of the Questions

Our first methodological step was aimed to assess the level of difficulty of the questions used throughout the experiment and baseline. To do this, we conducted a survey. It consisted of a HIT[3], published in AMT, with a survey embedded. In

---

[3] A HIT is an AMT an acronym for Human Intelligence Tasks. Generally, they are the result of the partition of large problem into small tasks, which any individual can work on through Amazon's platform.

addition to some demographic information the survey was composed of fifteen problem-solving questions. In particular each question has four answers but solely one was correct. From the four answers, participants could only choose one. AMT workers received $0.10 as an incentive to participate[4]. Additionally, we also tried to increase motivation by explicitly telling participants that in the end of the research the results would be shared with those who wished so. To avoid repetition, workers were prevented from participating more than once and hence, after delivering their survey, they were blocked and the results were filtered by response identification and IP address. Both questions and answers were randomized to avoid errors due to having a fixed dispositions. Lastly, we only considered results originated in OECD countries and, in particular, from a subset workers who had vocational/technical training or above. The aim was to obtain a set of questions evaluated by a similar crowd we would use in our experiment. In sum, we asked a crowd of AMT workers to evaluate the difficulty of a set of fifteen questions. From them we considered five to be used in the second methodological step.

3.2 Baseline

This methodological step was designed to create a baseline to compare other results against. For this stage we used the same survey used in the Experiment but without any bias. In this sense, participants' success in solving correctly the problems was solely dependent on their own skills. Through the baseline and the experiment was possible to establish a comparison between those who were socially influenced and those which were not. Two baselines were constructed through the answers of AMT workers and TBSE students. In both settings, participants who already performed the test were not allowed to participate again. The structure of the survey and the level of rewards were the same as the previous methodological step.

---

[4] The reward given was considered to be a reasonable incentive as, in accordance to Ipeirotis (2010), 90% of the HITs in AMT are subject to a payment of less than 10 cents.

3.3 Experiment

It consisted in administering two similar surveys in AMT worker community and to students from a top business school in Europe (TBSE). In contrast with the previous survey, at this stage participants performed a skill assessment in the form of an adaptive test and afterwards they were asked to solve the five problem-solving question. It is worth to note that we explicitly biased one answer in each of the five questions derived from the precedent step aiming to influence participants away from the correct answer. These issues are discussed in depth in the following section. Workers which took part in the previous survey were prevented from participating again. We ensured that every participant could answer the survey only once. In AMT the automatic code generator presented was used. Moreover, both the order of questions and the order of answers were randomized in order to avoid concerns of having the same arrangement of the two. Both TBSE students and AMT workers received the same intrinsic reward of having the possibility to access the results of the research in the end. Nevertheless, while TBSE students received no extrinsic reward, AMT workers received $0.10 for completing the HIT.

The analysis of the results of the baseline and the experiment provided insights about the existence of social influence in crowdsourcing platforms, as well as information about the circumstances in which it has more implications to participants' performance. To operationalize the attainment of the aforementioned results, the statistical analysis was performed through STATA 12 and JMP 10 software. The latter allowed us to understand the basic statistics while the former was used to compute the necessary regressions.

3.4 Survey Contents

The surveys used in the experiment and baseline was written in English and consisted of three parts. It was presented as an academic study conducted by university researchers with the aim of assessing the level of difficulty of a set of questions. In AMT, participants were informed about the structure of the survey, the

confidentiality agreement, compensation and the time available to complete the survey (15 minutes)[5]. After some general instructions and ask participants not to use calculators, the survey began. The same set of information, excluding compensation, were given to all participants (i.e. AMT participants and TBSE students).

The first part of the survey was designed to welcome participants and to obtain general demographic information.

The second part of the survey consisted in an adaptive numeracy skill assessment (Cokely et al., 2010). The aim was to ensure that we found participants with different levels of problem-solving skills as well as to test one of our hypotheses. The adaptive test format was used. This feature imply that the questions vary in terms of difficulty depending on the participants past success on answering previous answers. At least they have responded to two questions and in maximum three. Through this test it was possible to separate participants into four different levels of skills, ranging from 1 (low skill) to 4 (high skill). To assess the levels of problem-solving skills we use a freely available test (www.riskliteracy.org) named the Berlin Numeracy Test (BNT). This test is psychometric instrument that enable a quick assessment (on average 3 minutes) of statistical numeracy and risk literacy (Cokely et al., 2012). Moreover, it was tested in 21 studies (n=5336), across 15 countries with diverse samples, including Amazon Mechanical Turk. In the previous studies the BNT displayed a robust psychometric discriminability, in particular among highly educated individuals such as college students and graduates. In addition, this test was also correlated with other numeracy test and other measures of cognitive capabilities, such as Raven Advanced Matrices or Cognitive Reflection Tests. In sum, the aforementioned characteristics of the BNT demonstrate that the test was suitable to our purpose.

---

[5] Although we have made this request, some participants did not complied. Nevertheless, we felt that this would not be counterproductive to the current study and thus we kept the results arising therefrom.

The third part of the survey was composed of five problem-solving questions, similar in nature to those presented earlier. As stated in the previous section, the questions resulted from fifteen questions which we previously have submitted to AMT workers. As a consequence, the questions derived from that stage were characterized by having five distinctive levels of difficulty. The main reason to have different levels of difficulty stems from the fact that in previous face-to-face studies results have shown that social influence varies in function of the level of difficulty of the questions (Blake et al., 1957; Coleman et al., 1958). To choose which answers to bias we looked to the distribution of votes each answer received after the level of difficulty assessment. The rational chosen was to bias, in each question, the wrong answer that had received the highest number of votes. In this sense, in AMT, during the experiment participants were confronted with the following information: "the most popularly chosen answer is highlighted, and we also indicate how many other persons chose it in between brackets. E.g. [most popular answer with N votes]". In contrast, to account for the hypothesized impact of ties, in the survey distributed in TBSE we mentioned "the most popularly chosen answer is highlighted, and we also indicate how many students from TBSE chose it in between brackets. E.g. [most popular answer chosen by TBSE students with N votes]".

3.5 Data Analysis

3.5.1 Assessment of the Level of Difficulty of the Questions Data

In this stage, 171 AMT workers participated in this task. Answers from non-OECD countries were removed and we were left with 81 respondents. To ensure that the questions were correctly evaluated we considered solely the results of the set of highly educated individuals (vocational or technical training or above). 52 individuals complied with the criteria, in which 55% were female and 45% male. Individuals had a minimum age of 20 years old and a maximum of 65. The average was 37 years old. In what concerns the country of origin, 77% were from the US and 23% from other OECD countries (e.g. Australia, Canada, Italy, and Portugal). In addition, 81% were native English speakers. In terms of employment status, 31%

of the individuals were full-time employees and 18% were unemployed students. Among those which were employed, most of them work on Consulting Services (13%) or Health Care/Medical Services (13%) and the average annual income was between $20,000 and $29,999 US dollars. Most of the individuals had a college degree (58%), a master degree (19%) or vocational or technical training (19%). On average, participants took 9 minutes to complete the survey.

3.5.2 Baseline Survey Data

In the construction of the baseline were initially involved 451 individuals. 416 participants completed the survey which consists in a completion rate of 92%. There were 12 surveys which displayed the same IP address or response identification and hence were removed. From the 404 participants, 34% posted a comment and 56% shared their e-mail. Whereas AMT accounted for 91% of the participation, the remainder had origin on TBSE. In this sample 39% of the answers were originated in India, 37% in the US and the remaining 24% in 40 different countries. In terms of gender distribution, 60% were male and 40% were female. The younger individual which participated in the survey had 17 years old and the older 67. In general terms participants had, on average, 26 years old. In what concerns being a native English speaker, 47% were whereas 53% spoke other languages. The average annual income ranged between $5,000 and $9,999 US dollars. Similarly to the survey described previously, most of the participants (21%) had no income. Differently, in what concerns respondents' education level, they were on average college graduates (46%), though 26% were high schoolers or equivalent and 16% had vocational or technical training. The vast majority of respondents had a degree on Computer Science (14%), followed by Arts (11%) and Management (9%). While 31% were unemployed students, 30% were employed full-time. Participants were primarily working on the following industries: Information Services and Data Processing (7%); Health Care or Medical Services (6%); and Retail Commerce (6%). In what concerns participants' levels of problem-solving skills, 27% of the individuals had the lowest level of skills, followed by 34% of

medium-low level of skills and 18% and 21% of medium-high and high level of skills, respectively. On average participants took 11 minutes to finish their survey.

3.5.3 Experiment Survey Data

In overall terms, 891 individuals opened the survey and 791 finished it. This implies that the completion rate was around 89%. From those who answered it, we removed participations with the same IP address or response identification. We were left with 791, in which 38% commented the survey and 60% shared their e-mail in order to access the result in the end. It is worth to note that only 5% were from TBSE and the remainder from AMT. Approximately 55% of the participants were from India, 25% from the US and the remaining 25% from 54 different countries. Regarding the gender of the participants, 42% were female and the remainder were male. In what concerns the distribution of age, it ranged from 18 years old to 66 and, on average, participants had 28 years old. As the survey was worldwide distributed, one of our concerns was about participants' mother tongue. In this context, 42% were English native speakers. The average annual income was between $5,000 and $9,999 US dollars, though 18%, the largest percentage, had no income. Moreover, 48% of the participants were college graduates, 23% have a master's degree and 18% vocational or technical training. From those which have a college degree or above, 18% have a background on Computer Science while 11% had a degree on Management. Finally, 34% were employed full-time and 26% were unemployed students. The most significant professional occupations were on Computer and Electronics Manufacturing (11%), Adult Education (8%) and Health Care/Medical Services (6%). Considering the distribution of skills among the sample, 30% of the individuals had low skills (1), 33% medium-low skills, 16% medium-high and the remainder the highest level of skills. Lastly, participants needed, on average, 9 minutes to finish their survey.

At this stage is reasonable to assume that both baseline and experiment data are very similar in terms of the characteristics described above and thus provide a good basis for comparison.

3.6 Variables

3.6.1 Dependent Variable

The dependent variable used throughout this study was named Bias. It was represented by an informational cue posed upon the most voted wrong answer in the Experiment early described. The comparison of the results of the second (Experiment) and the third (Baseline) methodological step allow us to measure the effectiveness of the bias. Therefore, social influence is the difference between the number of votes a biased answer has received (Experiment) and the number of votes that same answer received without bias (Baseline).

In addition, this study sought also to deepen the knowledge about the variables that explain a greater or lesser propensity to be socially influenced. For this purpose regression analysis were made, having Bias as a dependent variable. This variable represents the choice of a wrong answer that contained the informational cue.

As the dependent variable can only take two values (to choose, or not to choose, the biased answer) we compute the regression through a *probit* model.

3.6.2 Independent Variables

In this study we considered several independent variables. Two of them depend on the personal characteristics of the individuals such as the numeracy skill level (*Skill_Level)*, as well as the time (*Test_Duration*) individuals need to complete their survey. Whereas the time was simple accounted in minutes, the skill level was the result of the adaptive test individuals previously performed.

The second set of independent variables were related with the context of the study. In this sense, we considered ties (*Ties*) between individuals and the test being biased (*Test_Biased*) to explain our dependent variable. Both are binary

variables but, while *Ties* distinguish TBSE from AMT individuals, *Test_Biased* differentiate the experiment from the baseline.

Lastly, cultural factors were weighted through the use of Hofstede's 1983 individualistic/collectivist indexes (*Hofstede_Idv*).

3.6.3 Control Variables

In this study we used control variables gender (*female*), age (*age*), annual income in US dollars (*income*), highest education level (*education*), field in which participants received the highest degree (*high_degree*), and employment status (*employment*).

Table 1. Summary of Variables

Control Variables	Independent Variables
Demographics female age income education high_degree employment	**Personal** skill_level test_duration **Contextual** ties test_biased **Cultural** hofstede_idv
Independent Variable	
bias	

IV RESULTS

To assert our main proposition, we start the presentation of the results by displaying two tables which statistically depict, on each trial, the difference between the number of votes the answers received in the experiment and in the baseline.

In order to understand which are the variables that explain being more or less permeable to pressures of the majorities in crowdsourcing we compute a regression on each of the five questions, having bias as the dependent variable. At this stage, a particular emphasis is given to question number 3. Due to the fact that it has an average difficulty within the scope of this study we present 6 regression models.

4.1 Comparison of the results of the Experiment and the Baseline

With the aim of asserting our main proposition about the existence of social influence in crowdsourcing platforms, in the tables 2 and 3 are displayed a comparison between the numbers of votes the biased answers received in the experiment, in contrast with the votes the same answer received on the baseline.

The results depict that when an informative cue is present, the number of votes a biased answer receive abruptly increase, on average terms, approximately 24% on AMT and 10% on TBSE. Additionally, TBSE displayed a less degree of proneness to become socially influenced, when compared to AMT workers. This result gives support to the first hypothesis (H1).

As mentioned previously, the questions have different levels of difficulty. While question 1 and 2 were easy, 3 and 4 were average and question 5 difficult. Although at a first glance the results appear to display a steady rise of the number of votes as long as the difficulty of the questions increase, we should take into account that the answers chosen to be biased were those which were wrong and received the largest amount of votes. As such, interpreting the results in such way could led to misleading outcomes.

Table 2. Comparison of the Results of the Experiment and Baseline on AMT

Source: AMT	Biased Answer (Question 1)	Biased Answer (Question 2)	Biased Answer (Question 3)	Biased Answer (Question 4)	Biased Answer (Question 5)
Experiment (% of votes)	32,00%	41,00%	47,00%	45,00%	62,00%
Baseline (% of votes)	8,00%	15,00%	28,00%	18,00%	39,00%
Δ (Exp.-Bas.)	24,00%	26,00%	19,00%	27,00%	23,00%

Table 3. Comparison of the Results of the Experiment and Baseline on TBSE

Source: TBSE	Biased Answer (Question 1)	Biased Answer (Question 2)	Biased Answer (Question 3)	Biased Answer (Question 4)	Biased Answer (Question 5)
Experiment (% of votes)	2,00%	9,00%	26,00%	26,00%	54,00%
Baseline (% of votes)	0,00%	3,00%	18,00%	18,00%	27,00%
Δ (Exp.-Bas.)	2,00%	6,00%	8,00%	8,00%	27,00%

4.2 Demographics

Table 4. Statistical Significance of Control Variables in Bias

Control Variables	Biased Answer to Question 1	Biased Answer to Question 2	Biased Answer to Question 3	Biased Answer to Question 4	Biased Answer to Question 5
Female	.	.	++	.	++
Age	.	.	.	.	.
Income	.	.	.	- -	.
Education	.	-	.	-	- - -
High_degree	.	.	.	.	++
Employment	++	.	.	.	.

Legend:
+++/- - - (positive/negative impact p<0.01)
++/- - (positive/negative impact p<0.05)
. (no statistically significant impact)

According to the results, in all of the five questions none of the control variables presented consistently explain the proneness to become socially influenced.

Of interest to the current research is however the fact that *education* starts to have more reliable statistical relevancy as long as the objective task become more difficult to solve. As such, this variable displays a negative coefficient of 0.126 (p<0.1) and 0.246 (p<0.01) in question number 4 and 5, respectively. Although none of the variables were, at this stage, directly implicated in our hypotheses testing, *education* was indirectly in accordance with our expectations since the experiment was composed of problem-solving questions and having a higher degree of education was expected to yield positive outcomes on answering correctly to the questions.

Similarly, the behavior of the variable *female* began to have some impact at the end of the experiment (question 3 and 5). The variable presented a positive coefficient of 0.183 (p<0.05) (question 3) and 0.229 (p<0.05) (question 5). Although we did not expect being more or less biased dependent on participants' gender, previous research on conformity supported the view that the greater the proportion of female respondents the greater the level of conformity (Bond & Smith, 1996).

4.3 Personal, Contextual and Cultural Variables

In what concerns the other independent variables related with personal, contextual and cultural factors, *skill_level*, *ties* and *test_biased* revealed to be consistent predictors of being more or less permeable to majorities' pressures.

Table 5. Statistical Significance of the Independent Variables on Bias

Independent Variables	Biased Answer to Question 1	Biased Answer to Question 2	Biased Answer to Question 3	Biased Answer to Question 4	Biased Answer to Question 5
Skill_level	.	- -	- - -	- - -	- - -
Test_dur	.	.	.	.	.
Ties	- - -	- - -	- - -	- - -	.
Test_biased	+++	+++	+++	+++	+++
Hofstede_idv	- - -	-	.	.	+

Legend:
+++/- - - (positive/negative impact $p<0.01$)
++/- - (positive/negative impact $p<0.05$)
+/- (positive/negative impact $p<0.1$)
. (no statistically significant impact)

Firstly, *skill_level* proved to be a significant predictor of *bias*. As such, when individuals display a good level of skills, they are less prone to follow the suggestions of the majorities. In the other way around, less problem-solving skills is a synonym of more susceptibility to engage in a replicative a behavior. In sum, this result supports H1.b.

It is noteworthy that, in opposition to what we have previously hypothesized in H1.a, *ties* have a strong negative impact on *bias*. In accordance, this mean that although individuals from a top European business school shared a common basis with the source of influence (hypothetically they were students of the same University), individuals were not engaged in a herding behavior. In contrast, according to the results, having ties with the source is a predictor of being less influenced. This result is analyzed in depth in the following chapter.

Differently, *test_biased* was regarded as a strong predictor of being biased. This means that when social influence is embedded in the setting, individuals were strongly impacted towards the biased answer. This findings are in accordance with table 2 and 3 and support our first hypotheses (H1).

Lastly, *hofstede_idv* did not prove to have a consistent significant impact on the dependent variable *bias*. In accordance with the results, participants from individualistic or collectivist countries did not show any systematic patterns of behavior according to the cultural inclination of his/her country.

Table 6. Statistical Significance of the Independent Variables on Bias (Question 3)

| Independent Variables	Model 1	Model 2	Model 3	Model 4	Model 5
Female	+	+	+	+	+
Age	.	.	.	.	.
Income	.	.	.	.	.
Education	.	.	.	.	.
High_degree	.	.	.	.	.
Employment	.	.	.	.	.
Skill_level	- - -	- - -	- - -	- - -	- - -
Test_dur		.	.	.	.
Ties			- - -	- - -	- - -
Test_biased				+++	+++
Hofstede_idv					-
Legend: **+++/- - -** (positive/negative impact p<0.01) **+/-** (positive/negative impact p<0.1) **.** (no statistically significant impact)					

Likewise, when looking specifically to question number three the same results are found. In the model number 6, the variables that have the higher coefficients are *test_biased*, *ties* and *skill level*, with values of, respectively, 0.544 (p<0.01), -0.536 (p<0.01) and -0.173 (p<0.01).

The following image depicts the relationships among the independent variables and *bias* in this same question. The thickness of the arrows illustrate the statistical impact a variable has on the dependent variable:

Figure 1. The Impact of the Independent Variables on Bias in Question 3 (Model 5)
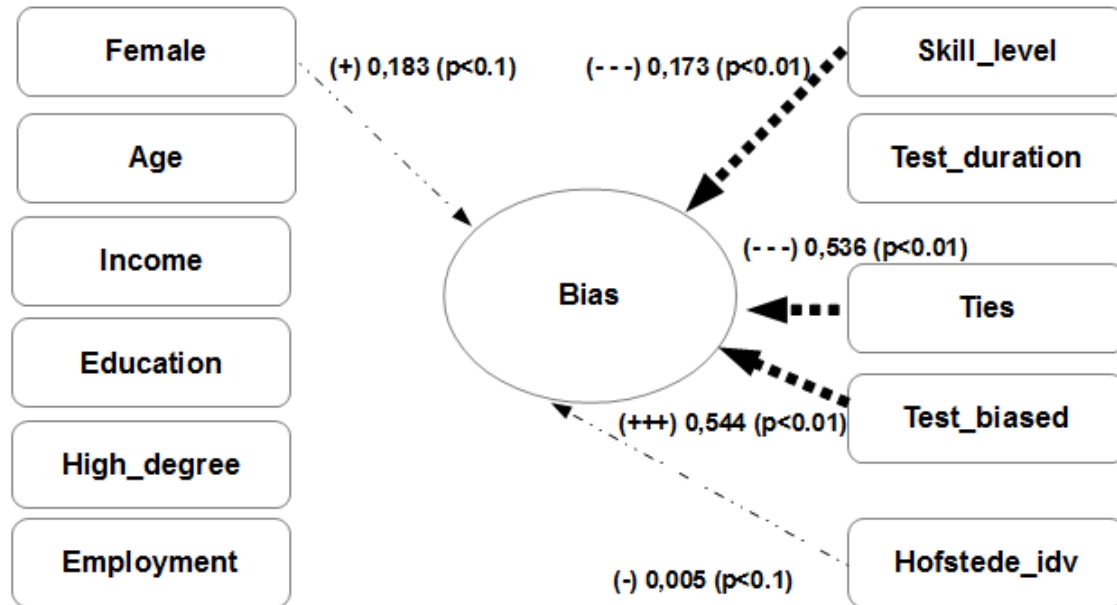
4.4 Analysis of the Comments

Throughout the study we were also able to collect information about what participants thought about their participation and the study itself. The two most noticeable types of comments were about the difficulty of the problems posed and the lack of time to solve them, as well as the fact that of the survey being interesting, fun and helpful.

Of relevancy to this research was the fact that some of the comments were about the bias. One participant mentioned that "the fact of having the most popular answer shaded is leading me to assume that's the right answer". Other stated that "in the last part it was really difficult for me to answer certain questions so I chose the best answer with higher votes". This is not more than one clue to indicate that the bias may have had the intended desideratum in some participants. In contrast, it should also be noted that other participants suspected of the influence posed. For instance, one participant stated that he "suspect that your "most popular answers" note was bogus. I would have been more inclined to believe it if not for the coin toss question. Surely not many people believe that a coin toss isn't 50-50!".

V DISCUSSION AND CONCLUSIONS

This study was designed to explore the existence of social influence in crowdsourcing platforms, in two different contexts: anonymous and social. Of specific relevance to the current research is that decision-making mechanisms such as voting, averaging, estimations, among others, are common in crowdsourcing initiatives. In this context, our point of depart was that the presence of informative cues, such as those mentioning what is the current number of votes a proposition has received, could undermine and – worse – pollute the goal of obtaining valuable results from crowds.

Additionally, we studied the circumstances in which social influence has more impact over the performance of individuals. To attain this purpose, we asked individuals to solve a set of problem-solving questions. We deliberately gave informational cues in order to introduce social pressures and thus influence participants towards a wrong answer.

Through the current investigation we found that individuals could be socially influenced even when they are asked to perform objective tasks for which there is one correct answer. Moreover, no matter the difficulty of the task they perform, individuals cede their independence towards the will of the majority when informational cues are present. This fact was supported by the evidence found in the difference between the numbers of votes the five biased answers received in the experiment in comparison with baseline. Additional corroboration was verified through the reliable statistical impact of the variable *test_biased* on *bias*. As such, when an answer has an informational cue we saw a steady rise of votes gravitating upon them. This result offered support to our first hypothesis (H1).

Additionally we tried to investigate if the existence of *homophilous* ties among participants of TBSE has any impact on the levels of conformity. In opposition to what was primarily expected, *ties* did not provide any source of additional pressure. As such, hypothesis H1a was not corroborated. From our point of view, two explanations might address this unforeseen result. Firstly, it was difficult to measure the degree of proximity between TBSE participants. The universe of

students of this school is differentiated in terms of country of origin and many of them are affiliated with the university for short periods of time. In result, this might have, for example, weakened the sense of belonging to the institution or to the students' community and thus reduce the degree of trustfulness over the source of influence. Secondly, we expect that the participants from this school were used to perform such type of problem-solving questions, as they were former or current Master degree students of Management or Economics. The latter argument gives some weight to the possibility of having participants solving themselves the questions or anticipating that the bias was always present upon the wrong answers.

More importantly, throughout this study we were interested in understanding what could be a stronghold against majorities' invectives. Empirically we provided evidence to support hypothesis H1.b, given that the level of skills a participant has is directly implicated on the level of conformity. Participants with a high level of skills are less prone to engage in replicative behavior. In contrast, the probability of a participant with low level of skills converge to the will of the majority is high.

Finally, we tried to understand if a country's cultural inclination towards individualism or collectivism could be a synonym of greater or lesser attachment to group decisions. The results were not very clear. On the one hand, this could be a symptom of having almost 80% of participation from only two different countries, India and United States of America. On the other hand, Hofstede's indexes, used to compute the results, date from the early eighties and this cultural dimension has shown a great level of convergence over the years (Hofstede, 1983). The latter assertion offers some explanation to our result: as countries increasingly display similarities in terms of the two cultural factors, no significant differences arise regarding the levels of conformity.

Finally, the question regarding whether the attained consensus through the pressures of majorities yield beneficial or detrimental outcomes should be considered. Social influence is a great source of shortcuts and consensus, which are indispensable to life in societies. It is however noteworthy that those pressures may result in a form of dominance. In this sense, they could undermine the behavior

of individuals and, above all, suppress independent contributions. One of the main benefits of crowdsourcing is articulating unused sources of knowledge that are valuable due to their specificity or aggregated contribution. As such, one might expect that when there are pressures to homogenize singular expressions, one key advantage of the crowdsourcing model fades away, invalidating the wisdom of crowd hypothesis.

5.1 Managerial Implications

Through this study we explored the existence of social influence in crowdsourcing. Accordingly, we found empirical evidence that when carrying out objective tasks, individuals are influenced by majorities. At first glance this result demonstrates that this phenomenon exists and informational cues about what others may think, in fact, condition the behavior of others. More striking, those clues led others to believe that the aggregated outcome of the crowd was in fact accurate. In this context, we believe that because crowdsourcing platforms are subject of an increasing use by companies, the first contribution of this investigation relies on the disclosure of this phenomenon.

Secondly, this study also provided evidence about what can be used to overcome this social dynamic. In this context, the cognitive skills of individuals proved to be an impregnable bulwark against the pressures of the majorities. In practical terms, this results suggest that firms necessarily need to have a high degree of responsibility in the crowd they choose to perform their tasks. More broadly, the results indicate that it is not exclusively a problem of seeking to broadcast a problem to a differentiated group of individuals, but also to ascertain that a specific crowd is the most suitable to provide independent and valuable results.

Finally it is worth mentioning that cultural variables not proved to be an element which constrain the behavior of individuals. Because of one the advantages of crowdsourcing is the ability to achieve contributions of locally scattered individuals, the origin of this same contribution, as far as it was possible

to understand, did not explains a greater propensity to be pressured by majorities. In a managerial point of view, this contribution underlies the benefits of calling upon a crowd locally distributed without incurring in a bias, as well as to demystify some theoretical contributions that regard social influence as a function of the level of a more or less collectivist/individualistic culture.

5.2 Limitations

The first concern this research faces is due to the fact our setting, using the AMT crowd, was static rather than dynamic. In crowdsourcing platforms other than AMT, individuals usually provide contributions throughout a consistent period of time and have different tools to express their decisions and opinions (comments; forums of discussion; direct contact with other participants, etc.). Differently, in our study such conditions were not met but they might prove to be important mechanisms to avoid pressures from the majorities.

Secondly, to more faithfully replicate the social influence phenomenon, the experiment could have contemplated the existence of minorities. As they are the result of dissonant perspectives, they engage individuals to reexamine their thoughts and behaviors (Moscovici & Personnaz, 1979). As such, their presence could have led to different results since the current experiment only provided information about what the majority thought was the correct answer. Additionally, accounting for personality traits could have led to a different approach as it could give an explanation for participants being more or less prone to follow the predictions of a crowd, independently of having higher or lower level of skills.

Thirdly, the use of Hofstede's individualistic indexes, as they dated from a study published in the eighties, are expectable to have changed throughout time. In fact, this argument gains even more weight if we take into account the fact that the individualistic dimension is the one who have shown the most dramatic shift and displaying a certain convergence among countries over time (Hofstede, 1983).

At last, the sample size of TBSE also offers some limitations to our study. Although several efforts were made to have access to more students, that step was

unfortunately not possible. As such, we were left with 7% of the participation being originated in TBSE. The existence of a larger sample would have provided a clear picture of the social phenomenon here subject to analysis. In addition, a larger TBSE sample would enable to overcome a critical challenge of this study that was measuring the effectiveness of the ties. For example, another sample from TBSE could be biased in the same form as AMT experiment was. As a result, some insights about the bias being more or less effective depending on the nearness of the source and the target of influence could be raised.

5.3 Future Research

In order to continue the research within the scope of social influence in crowdsourcing platforms, we consider that it would be important to understand this phenomenon in a dynamic context where participants could interact continuously and had at their disposal other mechanisms of social interaction, such as comments, online contacts, etc. (Centola, 2010).

Secondly, another area of future research concerns the possibility to provide evidence on whether social influence, or the existence of informative cues, damage or enrich crowdsourcing platforms in situations in which individuals have access to informative cues provided by minorities (Moscovici & Personnaz, 1979).

Finally, it would be relevant to understand and compare the level of social influence participants experience in ambiguous tasks, as well as to ascertain if other cognitive and personality characteristics explain the proneness of an individual being more or less influenced by a majority (Neulinger & Stein, 1971).

REFERENCES

Asch, Solomon E. (1955), "*Opinions and Social Pressure*", Scientific American, 193 (5): 31 – 35.

Alexa – the Web Information Company, "*Alexa Top 500 Global Sites*", http://www.alexa.com/topsites, accessed July 2012.

Benbya, Hind, and Van Alstyne, Marshall (2011), "*How to Find Answers Within Your Company*", MIT Sloan Management Review, 52 (2): 65 – 75.

Blake, Robert R., Helson, Harry, and Mouton, Jane S. (1956), "*The Generality of Conformity Behavior as a Function of Factual Anchorage, Difficulty of Task, and Amount of Social Pressure*", Journal of Personality, 25 (3): 294 – 305.

Bond, Rod, and Smith, Peter B. (1996), "*Culture and Conformity: A Meta-Analysis of Studies Using Asch's (1952b, 1956) Line Judgment Task*", Psychological Bulletin, 119 (1): 111 – 137.

Campbell, Jennifer D., and Fairey, Patricia J. (1989), "*Informational and Normative Routes to Conformity: The Effect of Faction Size as a Function of Norm Extremity and Attention to the Stimulus*", Journal of Personality and Social Psychology, 57 (3): 457 – 468.

Centola, Damon (2010), "*The Spread of Behavior in an Online Social Network*", Science, 329: 1194 – 1197.

Centola, Damon (2011), "*An Experimental Study of Homophily in the Adoption of Health Behavior*", Science, 224: 1269 – 1272.

Chesbrough, Henry W. (2003), "*The Era of Open Innovation*", MIT Sloan Management Review, 44 (3): 35 – 41.

Christakis, Nicholas A., and Fowler, James H. (2007), "*The Spread of Obesity in a Large Social Network over 32 Years*", The New England Journal of Medicine, 357 (4): 371 – 379.

Cialdini, Robert B., and Goldstein, Noah J. (2004), "*Social Influence: Compliance and Conformity*", Annual Review of Psychology, 55, 591 – 621.

Cialdini, Robert B., and Trost, Melanie R. (1998), "*Social Influence: social norms, conformity, and compliance*", In DT Gilbert, ST Fiske, G Lindzey (eds). *The Handbook of Social Psychology*, McGraw-Hill.

Cokely, Edward T., Galesic, Mirta, Schulz, Eric, Ghazal, Saima, and Garcia-Retamero, Rocio (2012), "*Measuring Risk Literacy: The Berlin Numeracy Test*", Judgment and Decision Making, 7 (1): 25 – 47.

Coleman, Janet F., Blake, Robert R., and Mouton, Jane S. (1957), "*Task Difficulty and Conformity Pressures*", The Journal of Abnormal and Social Psychology, 57 (1): 120 – 122.

Deutsch, Morton, and Gerard, Harold B. (1954), "*A Study of Normative and Informational Social Influences Upon Individual Judgment*", Research Center for Human Relations, 629 – 636.

Fowler, James H., and Christakis, Nicholas A. (2008), "*Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study*", BMJ, 337: 1 – 9.

Freedman, Jonathan L., and Fraser, Scott C. (1966). "*Compliance Without Pressure: The Foot-In-The-Door Technique*", Journal of Personality and Social Psychology, 4 (2): 195 – 202.

Gartner (2012), *Innocentive positions itself as a challenge marketplace*

Gerard, Harold B., Wilhelmy, Roland A., Conolley, Edward S. (1968). "*Conformity and Group Size*", Journal of Personality and Social Psychology, 8 (1): 79 – 82.

Granovetter, Mark S. (1973), "*The Strength of Weak Ties*", American Journal of Sociology, 78 (6): 1360 – 1380.

Hayek, Friedrich A. (1945), "*The Use of Knowledge in Society*", The American Economic Review, 35 (4): 519 – 530.

Hofstede, Geert (1983), "*National Cultures in Four Dimensions: A Research-based Theory of Cultural Differences among Nations*", International Studies of Management & Organization, 1 (2): 46 – 74.

Howe, Jeff (2006), "*The Rise of Crowdsourcing*", Wired Magazine, 14: 1 – 5.

Ipeirotis, Panagiotis G. (2010), "*Analyzing the Amazon Mechanical Turk Marketplace*", XRDS: Crossroads, The ACM Magazine for Students, 17 (2): 16 – 21.

Lakhani, Karim R., and Panetta, Jill A. (2007), "*The Principles of Distributed Innovation*", Innovations, 97 – 112.

Latané, Bibb (1981), "*The Psychology of Social Impact*", American Psychologist, 36 (4): 343 – 356.

Malone, Thomas W., Laubacher, Robert, and Dellarocas, Chrysanthos (2010), "*The Collective Intelligence Genome*", MIT Sloan Management Review, 51 (3): 21 – 31.

McDonnell, John, Domingo, Devin, and Gureckis Todd (2012), *Is Mechanical Turk the future of cognitive science research?* (Online) Available at: www.http://gureckislab.org/blog/?p=1297.

McPherson, Miller, Smith-Lovin, Lynn, and Cook, James M. (2001), "*Birds of a Feather: Homophily in Social Networks*", Annual Review of Sociology, 27: 415 – 444.

Milgram, Stanley (1973), "*The Perils of Obedience*", Harper's, 247 (1483): 62 – 77.

Morris, William N., Miller, Robert S., (1975) "*The Effects of Consensus-breaking and Consensus-preempting Partners on Reduction of Conformity*", Journal of Experimental Social Psychology, 11 (3), 215 – 223

Moscovici, Serge, and Personnaz, Bernard (1979), "*Minority Influence and Conversion Behavior in a Perceptual Task*", Journal of Experimental Social Psychology, 16, 270 – 282.

Nature (2005), *Internet encyclopaedias go head to head.*

Neulinger, John, and Stein, Morris I. (1971), "Personality Characteristics of Volunteer Subjects", Perceptual and Motor Skills, 32, 283 – 286.

Pisano, Gary (2006), "*Profiting from innovation and the intellectual property revolution*", Research Policy, 35, 1122 – 1130.

Postmes, Tom, Spears, Russell, and Lea, Martin (1998), "*Breaching or Building Social Boundaries? SIDE – Effects of Computer-Mediated Communication*", Communication Research, 25 (6): 689 – 715.

Rafaat, Ramsey M., Chater, Nick, and Frith, Chris (2009), "*Herding in Humans*", Trends in Cognitive Sciences, 13 (10): 420 – 428.

Schenk, Erik, and Guittard, Claude (2010), "*Towards a characterization of crowdsourcing practices*", Strasbourg: s.n., 1 – 20.

Tanford, Sarah, and Penrod, Steven (1984), "*Social Influence Model: A Formal Integration of Research on Majority and Minority Influence Process*", Psychological Bulletin, 95 (2): 189 – 225.

Travers, Jeffrey, and Milgram, Stanley (1969), "*An Experimental Study of the Small World Problem*", Sociometry, 32 (4): 425 – 443.

Triandis, Harry C., Bontempo, Robert, Villareal, Marcelo J., Asai, Masaaki, and Lucca, Nydia (1988), "*Individualism and Collectivism: Cross-Cultural Perspectives on Self-Ingroup Relationship*", Journal of Personality and Social Psychology, 54 (2): 323 – 338.

Villarroel, Andrei J. (2008), "*Open Source Corporate Strategy (OSCS): Unveiling the Firm's Open Sources of Competitive Advantage*" Ph.D. diss. EPFL.

Villarroel, Andrei J. (2011), "*Strategic Crowdsourcing: The Emergence of Online Distributed Innovation*", *Leading Open Innovation*. MIT Press.

Villarroel, Andrei J., and Reis, Filipa (2010), "*Intra-corporate crowdsourcing (ICC); leveraging upon rank and site marginality for innovation*" in Proceedings of Crowdconf 2010: The World's first conference on the future of distributed work, California, CA: San Francisco

Villarroel, Andrei J., and Reis, Filipa (2012), "*Crowdsourcing Corporate Innovation: Bridging Knowledge Silos of the Firm for Competitive Advantage*", 1 – 30.

Villarroel, Andrei J., and Tucci, Christopher L. (2009), "*Motivating firm-sponsored e-collective work*", MIT Sloan School of Management, Working Paper, 1 – 43.

von Hippel, Eric (1994), "*Sticky Information" and the Locus of Problem Solving: Implications for Innovation*", Management Science, 40 (4): 429 – 439.

von Hippel, Eric (2001), "*Innovation by User Communities: Learning from Open-Source Software*", MIT Sloan Management Review, 82 – 86.

von Hippel, Eric, and von Krogh, Georg (2003), "*Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science*", Organization Science, 14 (2): 209 – 223.

Watts, Duncan J., and Strogatz, Steven H. (1998), "*Collective dynamics of "small-world" networks*", Nature, 393 (4): 440 – 442.

Wellman, Barry (2001), "*Computer Networks As Social Networks*", Science, 293, 2031 – 2034.

LIST OF TABLES

LIST OF FIGURES

APPENDICES

Appendix 1: Probit Regressions

Dependent variable: Biased Answer to Question

Independent Variable	Biased Answer to Question 1	Biased Answer to Question 2	Biased Answer to Question 3	Biased Answer to Question 4	Biased Answer to Question 5
	coef/se	coef/se	coef/se	coef/se	coef/se
Female	0.015	0.141	0.183**	0.051	0.229**
	0.104	0.096	0.091	0.093	0.092
Age	-0.011	0.001	-0.008	-0.002	0.009
	0.007	0.006	0.006	0.006	0.006
Income	-0.015	-0.007	0.024	-0.042**	0.015
	0.021	0.019	0.018	0.019	0.019
Education	0.007	-0.138*	-0.056	-0.126*	-0.246***
	0.078	0.072	0.068	0.068	0.067
High_degree	0.000	0.006	0.002	0.005	0.010**
	0.006	0.005	0.005	0.005	0.005
Employment	0.100**	0.006	-0.004	0.047	0.012
	0.041	0.038	0.036	0.037	0.037
Skill_level	-0.008	-0.085**	-0.173***	-0.152***	-0.135***
	0.046	0.043	0.041	0.042	0.041
Test_duration	0.000	0.001	-0.000	0.000	0.000
	0.000	0.001	0.000	0.000	0.000
Ties	-1.793***	-1.074***	-0.536***	-0.387***	0.098
	0.439	0.254	0.192	0.197	0.182
Test_biased	0.962***	0.823***	0.544***	0.707***	0.706***
	0.126	0.109	0.097	0.101	0.095
Hofstede_idv	-0.010***	-0.005*	-0.005*	0.001	0.005*
	0.003	0.003	0.002	0.002	0.002

Appendix 2: Statistical Significance of the Independent Variables on Bias (Question 3)
Dependent Variable: Bias

Independent Variable	Model 1	Model 2	Model 3	Model 4	Model 5
	coef/se	coef/se	coef/se	coef/se	coef/se
Female	0,161*	0,161*	0,167*	0,155*	0,183*
	0,087	0,087	0,087	0,088	0,091
Age	-0,006	-0,006	-0,007	-0,009	-0,008
	0,006	0,006	0,006	0,006	0,006
Income	0,008	0,008	0,003	0,009	0,024
	0,016	0,016	0,016	0,016	0,018
Education	-0,019	-0,018	0,000	-0,028	-0,056
	0,064	0,064	0,064	0,066	0,068
High_degree	0,002	0,002	0,000	0,001	0,002
	0,005	0,005	0,005	0,005	0,005
Employment	0,037	0,037	0,017	0,012	-0,004
	0,034	0,034	0,035	0,035	0,036
Skill_Level	-0,165***	-0,165***	-0,159***	-0,169***	-0,173***
	0,040	0,040	0,040	0,040	0,041
Test_duration		0,000	0,000	0,000	0,000
		0,000	0,000	0,000	0,000
Ties			-0,486***	-0,461***	-0,536***
			0,185	0,187	0,192
Test_biased				0,528***	0,544***
				0,095	0,097
Hofstede_idv					-0,005*
					0,002

Appendix 3: Respondents' Country of Origin Relative Frequencies

Respondents' Country of Origin	Relative Frequency	Respondents' Country of Origin	Relative Frequency
Albania	0,26%	Morocco	0,17%
Algeria	0,17%	Netherlands	0,17%
Andorra	0,09%	New Zealand	0,09%
Australia	0,61%	Nigeria	0,17%
Austria	0,17%	Norway	0,09%
Bangladesh	0,43%	Oman	0,09%
Belgium	0,17%	Pakistan	0,52%
Bosnia and Herzegovina	0,09%	Peru	0,09%
Brazil	0,69%	Philippines	0,95%
Brunei Darussalam	0,09%	Poland	0,43%
Bulgaria	0,17%	Portugal	5,90%
Canada	1,13%	Qatar	0,09%
Chile	0,09%	Republic of Moldova	0,09%
China	0,17%	Romania	0,43%
Croatia	0,35%	Russian Federation	0,09%
Czech Republic	0,09%	Saudi Arabia	0,17%
Denmark	0,09%	Singapore	0,09%
Dominican Republic	0,09%	Slovenia	0,26%
Egypt	0,17%	South Africa	0,09%
France	0,26%	Spain	0,35%
Georgia	0,43%	Sri Lanka	0,78%
Germany	0,35%	Sweden	0,09%
Greece	0,17%	Switzerland	0,09%
Hungary	0,26%	Taiwan	0,09%
India	49,13%	Trinidad and Tobago	0,09%
Indonesia	0,09%	Turkey	0,09%
Ireland	0,17%	Uganda	0,17%
Israel	0,17%	Ukraine	0,09%
Italy	0,26%	United Arab Emirates	0,17%
Kenya	0,09%	United Kingdom of Great Britain and Northern Ireland	1,04%
Latvia	0,09%	United Republic of Tanzania	0,09%
Lesotho	0,09%	United States of America	28,91%
Lithuania	0,09%	Uruguay	0,09%
Malawi	0,09%	Venezuela	0,09%
Malaysia	0,26%	Viet Nam	0,17%
Mexico	0,17%	Yugoslavia	0,26%
		Zambia	0,09%

Appendix 4: Descriptive Statistics of the Independent Variables

Variable	N	Mean	Std. Dev.	Min.	Max.
Female	1152	1.4	0.49	1	2
Age	1152	28	9	17	67
Income	1152	$5000 - $9,999	3	I have no income	Over $150,000
Education	1152	College Graduate	1	No formal schooling	Doctoral Degree, Professional Degree (Phd, Md, Jd)
High_degree	906	Eng. Electrical	9	Anthropology	Sociology
Employment	1152	Employed part-time	2	Unemployment student	Retired
Skill_level	1152	2	1	1	4
Test_duration	1152	10	9	1	70
Ties	1152	0.07	0.24	0	1
Test_biased	1152	0.65	0.47	0	1
Hofstede_idv	1115	60	23	12	91