Lappeenranta
University of Technology

Juho Salminen

THE ROLE OF COLLECTIVE INTELLIGENCE IN CROWDSOURCING INNOVATIONS

Thesis for the degree of Doctor of Science (Technology) to be presented with due permission for public examination and criticism in the Auditorium of Ravintola Voitto, Salpausselänkatu 8, Lahti, Finland on the 4th of December, 2015, at noon.

Acta Universitatis
Lappeenrantaensis 671

Supervisors Professor Vesa Harmaakorpi
LUT Lahti
School of Business and Management
Lappeenranta University of Technology
Finland

Professor Tuomo Uotila
LUT Lahti
School of Business and Management
Lappeenranta University of Technology
Finland

Reviewers Professor Pekka Kess
Department of Industrial Engineering and Management
University of Oulu
Finland

PhD Kaisa Still
VTT Technical Research Centre of Finland
Oulu
Finland

Opponent Professor Pekka Kess
Department of Industrial Engineering and Management
University of Oulu
Finland

ISBN 978-952-265-875-3
ISBN 978-952-265-876-0 (PDF)
ISSN-L 1456-4491
ISSN 1456-4491

Lappeenrannan teknillinen yliopisto
Yliopistopaino 2015

Abstract

Juho Salminen
The role of collective intelligence in crowdsourcing innovations
Lappeenranta 2015
195 pages
Acta Universitatis Lappeenrantaensis 671
Diss. Lappeenranta University of Technology
ISBN 978-952-265-875-3, ISBN 978-952-265-876-0 (PDF), ISSN-L 1456-4491, ISSN 1456-4491

Social insects are known for their ability to display swarm intelligence, where the cognitive capabilities of the collective surpass those of the individuals forming it by orders of magnitude. The rise of crowdsourcing in recent years has sparked speculation as to whether something similar might be taking place on crowdsourcing sites, where hundreds or thousands of people interact with each other. The phenomenon has been dubbed collective intelligence. This thesis focuses on exploring the role of collective intelligence in crowdsourcing innovations. The task is approached through three research questions: 1) what is collective intelligence; 2) how is collective intelligence manifested in websites involved in crowdsourcing innovation; and 3) how important is collective intelligence for the functioning of the crowdsourcing sites. After developing a theoretical framework for collective intelligence, a multiple case study is conducted using an ethnographic data collection approach for the most part. A variety of qualitative, quantitative and simulation modelling methods are used to analyse the complex phenomenon from several theoretical viewpoints or 'lenses'. Two possible manifestations of collective intelligence are identified: discussion, typical of web forums; and the wisdom of crowds in evaluating crowd submissions to websites. However, neither of these appears to be specific to crowdsourcing or critical for the functioning of the sites. Collective intelligence appears to play only a minor role in the cases investigated here. In addition, this thesis shows that feedback loops, which are found in all the cases investigated, reduce the accuracy of the crowd's evaluations when a count of votes is used for aggregation.

Keywords: collective intelligence, wisdom of crowds, crowdsourcing, innovation process, distributed cognition

Acknowledgements

In 2007 I was in exchange in Eindhoven, Netherlands. In the university library I stumbled upon a National Geographic Magazine featuring an article about swarm intelligence of honey bees. The article described various ways how collective intelligence is used in nature and business applications, and gave me an inspiration to learn more about the phenomenon. The results are presented in this thesis, but perhaps more important is the learning journey that lead to them. Along the way many people have offered me invaluable support.

First, I would like to thank my supervisor Professor Vesa Harmaakorpi for his guidance and support. Without him this project would not have been possible. I also wish to thank Professor Tuomo Uotila for constructive advice and feedback in various phases of the research project.

I wish to acknowledge the reviewers, Professor Pekka Kess from University of Oulu and Ph.D. Kaisa Still from VTT, for their valuable comments that helped in improving the quality of this manuscript. Professor Kess also agreed to be my opponent, for which I am grateful.

I would also like to thank my all my colleagues at LUT Lahti for their collaboration during various research projects and for the many useful discussions and encouragement along the journey of writing this dissertation. More importantly, they made coming to work a joy. Special thanks goes to Mrs. Raija Tonteri and Mrs. Hilkka Laakso for helping me with practicalities during both the dissertation work and other projects. I am also grateful for all the help and advice I received during my research exchange in the Stanford Center for Design Research, especially Ph.D. Tanja Aitamurto, with whom the collaboration has continued ever since.

I sincerely appreciate the financial support I received from Päijät-Häme Regional Fund of the Finnish Cultural Foundation and Vitako Ry. Their funding gave me the freedom to concentrate on the dissertation work.

Finally, I would like to express my gratitude for my family and friends. If I have learned anything, it is that, after all, the important things in life happen outside academia. I am fortunate to have you in my life.

A major part of this dissertation was written in various ‘coffices’, including Teerenpeli, Oskarin Piha, Tryffdeli, Robert’s Coffee and Café Venetia.

Juho Salminen
October 2015
Lahti, Finland

Contents

List of publications

This thesis contains material from the following papers. The rights have been granted by the publishers to include the material in this dissertation.

I. Salminen, J. (2012). Collective intelligence in humans: A literature review. *Collective Intelligence 2012*. Boston.
II. Salminen, J. (2013). Collective intelligence on a crowdsourcing site. *The Global Brain Institute Working Paper*.
III. Salminen, J. (2013). Crowdsourcing innovations as a search process. In Melkas, H. and Buur, J., eds., *Proceedings of the Participatory Innovation Conference*, pp. 320–323. Lahti.
IV. Salminen, J. (2014). Wisdom of crowds in practice. *Collective Intelligence 2014*. Boston.

Author's contribution

I am the principal author and investigator in all above papers.

1 Introduction

A swarm of honeybees is resting on a tree branch. A few hundred scouts are searching the surroundings in the hopes of finding a suitable new nest site: a closed dry cavity, big enough to host the colony, with a single small entrance. It is a matter of life and death. Failure to find one will mean the destruction of the entire honeybee colony within a couple of days, whereas a poor choice of nest site will expose the colony to weather and predators. Before long, one of the scouts finds a suitable looking tree cavity. It investigates the site thoroughly to determine its quality, a challenging task for the bee's pinhead-sized brain. After forming an opinion the bee returns to the swarm to announce the finding to other scouts with a waggle dance. A few fellow bees following the dance get interested, and, following the instructions on direction and distance conveyed by the waggle dance, fly to investigate the potential nest site. If they like it, they too return to the swarm to advertise the site to other scouts, recruiting ever more traffic to the site. The same process is repeated for sites discovered by other scouts, which creates a competition between the options. After a few hours, bees returning from one of the potential nest sites have changed their behaviour: the large number of bees at the nest site has led them to conclude that the decision on the nest site has been made. They signal the resting bees to prepare for take-off. Once the swarm is airborne, the scouts lead the swarm to their new home. Amazingly, the selected site is usually the best one available in the surroundings. Although each bee has limited cognitive capabilities and most scouts see only one of the options, the swarm as a whole is able to arrive at the correct decision. The phenomenon is an example of swarm intelligence: the cognitive capabilities of the swarm are orders of magnitude greater than the capabilities of its constituent parts.

The selection of nest sites is a crucially important decision for social insect colonies. Typically, the founding female makes this decision individually, but in some species of ants, bees and wasps the decision about the nest site is made collectively. Biologists have identified striking similarities between nest-site selection processes across different species, despite the fact that they have all evolved the required social behaviours independently of each other (Visscher 2007). Separate insect species have converged to similar solutions. These nest-site selection processes have also been found to be scalable and to fit well with the needs of different colony sizes (Franks et al. 2006). In particular, the nest-site selection process of honeybees has been studied thoroughly and is among the most complex known examples of self-organising group decision making in social insects (e.g., Seeley and Buhrman 1999; Seeley et al. 2006; Passino et al. 2008; Visscher 2007). Honeybees use the so-called weighted additive strategy in their decision making, which is cognitively demanding (Seeley and Buhrman 1999; Visscher 2007). In weighted additive strategy, the relevant attributes of each compared alternative are evaluated and given weights depending on their relative importance. The weighted evaluations are then combined and the best overall option is selected. According to simulation models, natural selection has tuned the parameters of this process close to the optimum compromise between the speed and accuracy of the decision-making process (Passino and Seeley 2005). The swarm's attention turns quickly to better quality sites as poorer quality sites are dropped from consideration (Passino et al. 2007). In this way, the resources of the swarm are directed to the evaluation of the best candidate sites ensuring that the probability of a bad decision remains low (Passino and Seeley 2005). During the process, individual bees rely only on local information; direct comparison of the nest sites is not necessary. All the available information is taken into account but none of the bees has to hold all that information. The bees even use an exponential scale in the evaluation, which amplifies the perceived differences of the nest sites (Seeley et al. 2006). Even though individual bees follow simple rules of thumb and use only locally available information, the self-organising system is able to integrate the information in a meaningful and useful way (Conradt and Roper 2005; Visscher 2007).

An intriguing question is whether something similar might be going on in the interactions of humans: could a group of humans have cognitive capabilities that are orders of magnitude greater than those of individual humans? After all, honeybees, ants and other social insects are not the only species on our planet facing challenging cognitive tasks. Economic development has remained a fundamental concern for human beings for centuries. The added challenges of approaching (or having already passed) planetary boundaries (Steffen et al. 2015) and climate change (IPCC 2014) do not make the task of improving human conditions any easier. Fundamentally, there are only two ways to increase economic output: by increasing the inputs or by figuring out ways to get more output from the inputs (Rosenberg 2003). The economic impact and importance of innovation has been accepted at least since the 1950s, when Abramovitz discovered that increases in inputs accounted for only about 15% of the growth of the United States economy between 1870 and 1950 (Abramovitz 1956). If anything, the importance of innovation has only increased since. A survey of Chief Executive Officers (CEOs) sitting at the top of 1,201 organisations in 69 countries identified innovation as one of three clear strategic focal points (PWC 2011). Although the CEOs had reported better penetration of their existing markets as the single best opportunity for growth since 2007, innovation now appears to be equally important for them. Innovation remained a priority in a 2012 survey (PWC 2012). Innovation is also of interest to people outside business organisations. The European Commission has placed innovation at the centre of its Europe 2020 strategy (European Commission 2010).

At the same time, the rise of the Internet has enabled new forms of collaboration to emerge, prompting speculation on the existence of collective intelligence of humans. The concept collective intelligence is still fuzzy and allows for many different interpretations, such as the comparison with the general intelligence factor of individuals (Woolley et al. 2010), the wisdom of crowds (Surowiecki 2005) and the swarm intelligence of social insects (Bonabeau 1999). Despite the undeveloped state of the concept, or perhaps because of it, academic interest in collective intelligence has exploded in recent years. Figure 1.1 shows an increasing trend in the number of articles published on the topic per year, as listed in the Web of Science Core Collection.

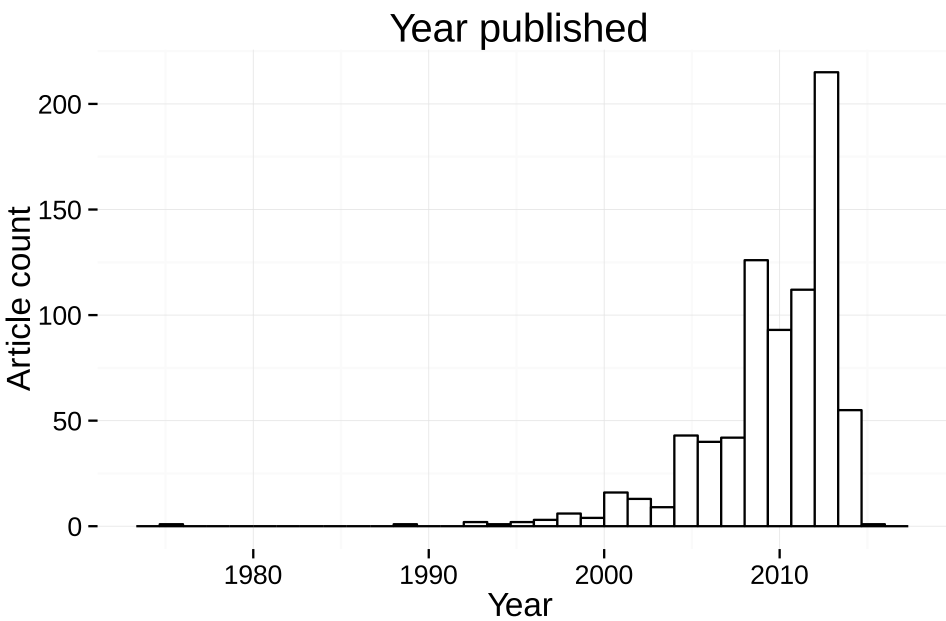


Figure 1.1: Number of published articles discussing collective intelligence in Web of Science Core Collection (Web of Science 2014). In total 785 records were found with search term “collective intelligence”.

Figure 1.2 lists the most popular keywords associated with the articles discussing collective intelligence. Although collective intelligence is clearly the most popular keyword, the others that follow are more interesting as they reveal connections between different concepts. The most popular keywords suggest that collective intelligence is related to phenomena residing on the Internet, such as interactive websites (web 2.0), crowdsourcing, social media and social networks. All these terms suggest the facilitation of interactions between large numbers of people over the Internet, but the connection to crowdsourcing is especially interesting. Crowdsourcing refers to the outsourcing of a task usually carried out by an organisation to an undefined crowd via an open call (Howe 2006). It is an approach that companies and other organisations can use to tap into the skills and knowledge of the masses. Among many other applications, crowdsourcing tasks related to the creation of innovation have gained particular attention. Crowdsourcing is still very much in the experimental state; although many organisations already rely on it, clear best practices have not yet emerged.

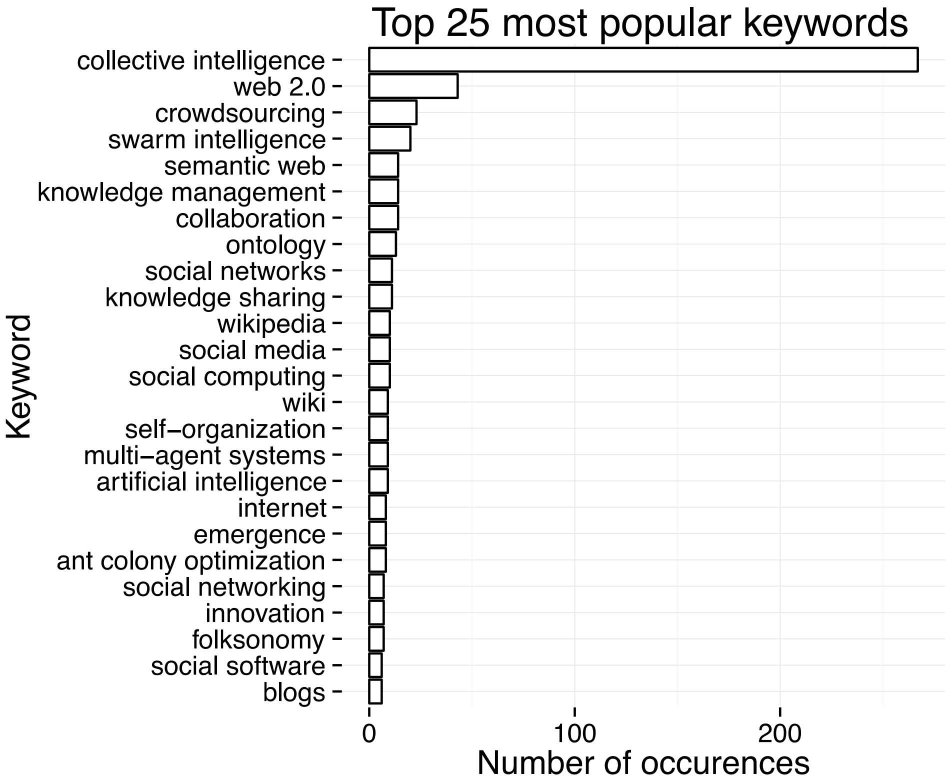


Figure 1.2: The most popular author keywords on articles on collective intelligence in Web of Science Core Collection (2014). In total 785 records were found with search term "collective intelligence".

An assumption seems to be that crowdsourcing supports, uses or benefits from collective intelligence. The idea that crowdsourcing might be one form of 'universal, distributed intelligence arising from the collaboration and competition of many individuals' (Levy 1997) is certainly appealing. If crowdsourcing indeed facilitates cooperation between humans similar to the swarm intelligence of social insects, we could expect dramatic improvements in our collective ability to create innovations. Such improvements could lead to significant reductions in the failure rates currently observed in new product development, which tend to be around 40%, depending on the industry (Castellion and Markham 2012). Speculation abounds on the significance of new forms of collaboration and interaction facilitated by the Internet. Global brain has been suggested as a metaphor for emerging, collectively intelligent networks formed by people, computers and the communication links connecting them (Heylighen 2011). The MIT Center for Collective Intelligence focuses on studying how people and computers can be connected so that collectively they act more intelligently than any person, group, or computer has ever done before (MIT Center for Collective Intelligence 2015).

Unfortunately, evidence on the relationship between crowdsourcing and collective intelligence is still lacking. The study of the swarm intelligence of social insects has revealed many interesting details on the collective properties of swarms of simple agents and has led to the development of dozens of practical applications for optimisation, robotics, data mining and classification (e.g., Dorigo et al. 2000; Mondada et al. 2004; Sousa et al. 2004). Research on human collective intelligence, however, is trailing behind. The problem is that we simply do not know whether something best described as collective intelligence actually emerges on crowdsourcing sites. For instance, should practitioners of crowdsourcing aim for collective intelligence and, if so, how can it be done? What is the role of the wisdom of crowd effect in crowdsourcing applications? Does something similar to the swarm intelligence of social insects take place on crowdsourcing sites when humans are interacting with each other? How important is collective intelligence, whatever it might mean, to the performance of crowdsourcing sites? While crowdsourcing is gaining popularity and has even been used to support political decision making (Aitamurto and Landemore 2013; Landemore 2014), we should be certain it is actually a good idea and at the very least that it does not promote collective stupidity or ‘madness of crowds’ (Mackay 1841). We might risk either missing a great opportunity or getting carried away by a hype bubble. Separating the wheat from the chaff takes effort.

Table 1.1: Examples of collective level cognitive capabilities that vastly exceed capabilities at the individual level.

Example	**Individual level**	**Collective level**
Nest-site selection of honey bees	Individuals decide whether they like a candidate site	Swarm selects the best available nest site in the environment
Foraging of social insects	Individuals search for food sources, collect food and advertise food sources to others	Colony optimises foraging among different food sources
Brain	Individual neurons integrate and send signals	Consciousness
Crowdsourcing	Individuals interact on a website	Collective intelligence?

Some examples of system-level cognitive capabilities exceeding local-level capabilities by orders of magnitude are listed in Table 1.1. In nest-site selection, relatively limited numbers of insects search for and evaluate nest sites. They make errors in evaluations and usually get to see only one of the options. Nevertheless, the swarm is able to arrive at the best decision most of the time (Visscher 2007). Many species of social insects have evolved surprisingly effective systems for foraging. Individual insects act only on local information using very limited cognitive capabilities; however, at the colony level, the exploitation of different food sources is optimised (Camazine et al 2001). The brain consists of billions of neurons. Each of them responds to incoming electro-chemical signals by sending electro-chemical signals in turn. The complex interactions of neurons give rise to consciousness (Thagard et al. 2014). In crowdsourcing, hundreds or thousands of participants interact with each other. It is possible that some kind of higher-level cognitive capabilities could emerge from such interactions. The first three items on the list are known examples, but the last is mostly speculation. This study aims to shed light on higher-level cognitive capabilities possibly emerging on crowdsourcing sites.

1.1 Purpose of the study research design

The purpose of the study is to look for evidence of collective intelligence on crowdsourcing sites. Crowdsourcing applications are found in very different fields, ranging from small routine tasks to photography and design services, science, public policy making (Howe 2006, Doan et al. 2011, Aitamurto and Landemore 2013), but here the focus is specifically on sites that use crowdsourcing for the creation of innovations. The increasingly popular (Doan et al. 2011) use of crowdsourcing as a part of the innovation process promises benefits by allowing more people to participate in the creation of innovations. The assumption is that the innovation process will benefit from the participation of more people as they bring in new knowledge, skills and diverse viewpoints (Terwiesch and Xu 2008). Even though collective intelligence is often mentioned in connection with crowdsourcing (e.g., Bonabeau 2009; Malone et al. 2010; Brabham 2008b; Sullivan 2010), it is not clear whether it is in fact a relevant concept to describe what happens at crowdsourcing sites. The research questions this study seeks to answer are:

1. What is collective intelligence?
2. How is collective intelligence manifested in websites involved in crowdsourcing innovation?
3. How important is collective intelligence for the functioning of crowdsourcing sites?

To answer the questions, a systematic literature review on collective intelligence was conducted, culminating in the development of a theoretical framework to guide the research project. The literature review reveals three levels of abstraction in the discussion about collective intelligence in humans: the micro level, the macro level and the level of emergence. This conceptual framework is used to organise relevant themes and to identify directions for further research. Then, guided by the framework, an in-depth investigation of three crowdsourcing sites focused on innovation was conducted. Inspired by a classic treatise on distributed cognition, *Cognition in the Wild* (Hutchins 1996), and 30-day challenges (e.g., Spurlock 2004), I visited the websites of OpenIDEO, Quirky and Threadless on at least 30 days and collected data as a participant observer. Following the example set in the *Essence of Decision* (Allison and Zelikov 1999), several theoretical frameworks were used to analyse the ethnographically-collected data. Multiple methods, including qualitative and quantitative approaches and simulation modelling, were used to break down the complexities of the cases.

1.2 Contribution

As a main result of the analyses, two candidates for collectively intelligent phenomena on the crowdsourcing sites could be identified: 1) virtual discussions hosted on the websites; and 2) the wisdom of crowds in evaluating content submitted to the websites. A correlation between crowd evaluations and expert decisions exists, but it is not strong enough to be relied upon alone in decision making. Further investigation with simulation models revealed that the feedback loops found on all the studied sites could decrease the accuracy of crowd evaluations, especially if the evaluations were aggregated using simple vote counts, as is often the case.

1.3 Scope and limitations

The scope of this thesis intersects the fields of innovation management, crowdsourcing and collective intelligence. Innovation management forms the background for the study, whereas the research project is conducted in the context of crowdsourcing. The scope of the study is limited to crowdsourcing innovations, because there are many different applications of crowdsourcing from photography to micro-tasks. Such different applications are likely not comparable in terms of their relationship to collective intelligence. The main focus and contributions are on the emerging field of collective intelligence, as defined in the systematic literature review in Chapter 3.

Three similar cases where crowdsourcing is used as a part of innovation or product development processes are investigated in detail in a search for evidence of collective intelligence. The results are not generalizable to general population of crowdsourcing sites, as might be the case for the results obtained from studies on larger, randomly selected samples. Instead, the findings are generalizable to analogical cases, but not necessarily applicable to all crowdsourcing applications, even less to innovation management in general.

1.4 Structure

The thesis is structured as follows. Chapter 2 describes the philosophical orientation, researcher choices, and the methodology relied upon during the research. In Chapter 3 the relevant literature on innovation, crowdsourcing and collective intelligence is reviewed. The main focus is on collective intelligence, for which the results of a systematic literature review are reported. The chapter culminates in the development of a theoretical framework, which is used to direct the rest of the research. Three investigated cases are presented in Chapter 4. Chapter 5 is devoted to presenting the results of cross-case analyses. Several analytical frameworks are used to investigate the operation of the crowdsourcing sites in order to account for the different interpretations on how collective intelligence might manifest itself. In Chapter 6 the meaning of the results is reflected upon and the contributions of the thesis are positioned within the context of existing research. Final conclusions on the thesis are provided in Chapter 7.

1.5 Key concepts

Collective intelligence: Two or more individual humans independently, or at least partially independently, acquire information and these different packages of information are combined and processed through social interaction, which provides a solution to a cognitive problem in a way that cannot be implemented by isolated individuals (adapted from Krause et al 2009).

Crowdsourcing: Crowdsourcing refers to outsourcing of a task usually carried out by an organisation to an undefined crowd via an open call (Howe 2006). In other words, a crowdsourcing system enlists a crowd of humans to help solve a problem defined by the system owners (Doan et al. 2011).

Distributed cognition: A system, where cognitive labour is distributed. Individual agents form only a part of the system, and other parts of the systems, such as technical devices, can also do important cognitive work. For instance, a pen and paper can be used to store information. (Hutchins 1996).

Ethnography: An open-ended research practice that is based on participant observation. It focuses on the local and particularistic knowledge of the meanings, practices and artefacts of a particular social group (Kozinets 2002).

Innovation: The result of implementing a solution that addresses a problem or need, where either problem, solution, or their combination is new.

Innovation process: A description of tasks that usually need to be carried out to create an innovation. Innovation process is about identifying a problem, searching for a solution and putting the solution into practice.

Netnography: Ethnography conducted on the internet.

Swarm intelligence: Collective, largely self-organised behaviour emerging from swarms of social insects, where the cognitive capabilities of the swarm are orders of magnitude greater than the capabilities of its constituent parts (Bonabeau and Meyer 2001).

Wisdom of crowds: A phenomenon, where under the right circumstances, the aggregated judgment of a crowd can be closer to the truth than that of the best individuals in the crowd (Surowiecki 2005). For example, the average of several individuals' estimates can be accurate even if individual estimations are not.

2 Methods

In this chapter the methodology used in the study are presented. The chapter begins with a discussion on the philosophical choices and assumptions. Then the methodological choices regarding the research approach, a variation of ethnography, are justified. Overall research design is explained. The focus is on data collection techniques and general guidelines on qualitative and quantitative data analysis and simulation modelling to be followed. The chapter ends with a note on research ethics.

2.1 Philosophy of science

Philosophically, the orientation of this study is scientific realism. Scientific realism is committed to the view that there is an external reality that is separate from our descriptions of it; natural and social sciences can and should apply the same kinds of approaches to the collection and analysis of data. Godfrey-Smith (2002) defines scientific realism as the naturalisation of common sense realism:

> We all inhabit a common reality, which has a structure that exists independently of what people think and say about it, except insofar as reality is comprised of thoughts, theories, and other symbols, and except insofar as reality is dependent on thoughts, theories, and other symbols in ways that might be uncovered by science.

One reasonable aim of science is to produce accurate descriptions of what reality is like. Scientific realism is committed to the existence of a world that is independent of the mind which the sciences investigate. Scientific claims are interpreted literally and theoretical statements taken at their face value. Claims about both observable and unobservable concepts, properties and relationships are assumed to be either true or false. These literally-interpreted theoretical claims are knowledge about the mind-independent reality. The best scientific theories are thus able to give approximately true descriptions of the world (Chakravartty 2014).

Scientific realism is compatible with both qualitative and quantitative approaches. Most importantly, it is compatible and consistent with itself. For example, if all knowledge is arrived at through observation of facts, as positivism claims, then how can we know about positivism? What observations show that observations are the only source of knowledge, especially when counterexamples abound including theoretical physics and simulation models on many fields? On the other hand, the distinction between people, social phenomena and natural objects, which interpretivism (Bryman and Bell 2007) holds important, is necessarily arbitrary. Molecules making up the nerve cells are clearly natural, and so probably are the neurons themselves and the networks they form. Automatic information processing performed by these networks on visual signals appears to be a natural phenomenon. It leads to the recognition of a familiar face, for example, and an emotion triggered by that recognition. The brain uses similar automatic processes to create a subjective interpretation about the situation in predictably biased ways (e.g., Kahneman 2010). When questioned, people can report their interpretations, which are still affected by the group dynamics of the situation. Group dynamics can be successfully investigated using the natural sciences approach (Sterman 2000). Taking all this into account, where exactly are the boundaries of the subjective interpretation that make people different from natural phenomena? In this context, scientific realism saves one from a lot of worry about philosophical questions. There exists a reality about which we can know something. Scientific realism is in line with common sense and, through its belief in the approximate correctness of science, it can update itself according to new scientific findings. Other philosophical orientations tend to position themselves as immune to criticism from the sciences, which makes them look suspiciously like dogma. A good philosophy of science should be flexible in relation to new knowledge, in a similar way to what the fourteenth Dalai Lama has said about religion: 'If science proves some belief of Buddhism wrong, then Buddhism will have to change' (Gyatso 2005).

2.2 Studying complex social systems

In this study, the emergence of collective intelligence is framed as a complex adaptive system. Complex adaptive systems are dynamic nonlinear systems that can display self-organising behaviour; that is, they can create order solely on the basis of the interactions between the system components and without external coordination. Theories of complex adaptive systems were originally developed in physics, where very simple systems were found to be able to display surprisingly complex behaviour. These ideas have migrated to the social sciences, where it has been suggested that the study of how complex global patterns emerge from local interactions could have a significant impact on the field (Lansing 2003).

Agar (2004) argues that in order to understand complex adaptive systems, ethnography should be used due to the compatibility of assumptions and objectives. As it is not known what will emerge from the complex interactions and how (finding out is the goal of studying the complex system in the first place), the methodological issues on what data to collect and how only come up during the research project and cannot be completely planned in advance. In contrast to more traditional social science research approaches, the variables are not decided in advance. Instead, meaningful connections and patterns are noted during the research. The goal is to build explanations that include the unexpected things that are noted, not to concentrate solely on what one is supposed to notice.

Complex adaptive systems consist of many agents interacting with each other in complicated patterns: ethnographers describe complicated patterns with many links among many objects. Traditional social research approaches have their place in understanding how the social world works by building on ethnographic findings. Güney (2010) provides further justification by describing the theory-method fit between complex adaptive systems and ethnographic research in more detail. The theory of complex adaptive systems states that it is unrealistic to assume fixed relationships between agents in the system due to their constant shifting in reaction to the environment and each other. Studying such systems requires a methodology capable of capturing the dynamic behaviour of social agents. The fundamental assumption in ethnographic research is that social reality emerges out of the meanings that the participants create in local interactions. Participant observation, open-ended interviews and document analysis are necessary tools in capturing the emergent process. This kind of ethnographic research is important because of the need for evidence about participants' understanding about why they are doing what they are doing in the social system (Güney 2010).

Ethnographic approaches have been successfully used to study distributed cognitive systems, such as the navigation systems of an aircraft carrier (Hutchins 1995) and airline cockpits (Hutchins 1995; Hutchins and Klausen 1995). Distributed cognitive systems are discussed in more detail in Chapters 3.3.7 and 5.5, but for now it is sufficient to know that they are socially distributed and embodied. The study of such systems cannot be separated from the study of culture because the agents live in complex cultural environments. The theory states that cognitive activities use both internal and external resources, and that the meaning of activities depends on the context. Therefore, there is no substitute for technical expertise in this domain. As a result, participant observation is an invaluable part of studying distributed cognitive systems.

2.3 Ethnography and netnography

Ethnography is an open-ended research practice that is based on participant observation. It focuses on the local and particularistic knowledge of the meanings, practices and artefacts of a particular social group (Kozinets 2002). This has consequences for research design: instead of planning everything beforehand, methodological issues are expected to arise during the research as it develops in unforeseen ways. This flexibility is one of ethnography's greatest strengths. Although no two ethnographic studies are ever carried out in the same way, the research process usually involves certain phases: induction into the culture, gathering and analysing data, ensuring trustworthy interpretation and feedback from members of the social group (Kozinets 2002).

In this study, the ethnographic research is carried out on the web, an approach sometimes called netnography (Kozinets 2010). Netnography is a qualitative research methodology that adapts ethnographic research techniques to study communities emerging on the internet (Kozinets 2002). It uses publicly available information in online forums as the main data source and can therefore be conducted in unobtrusive manner. Like ethnography, netnography is inherently flexible and adaptable. The largest deviation from traditional ethnography is the way the data is collected. The most important data sources in netnography are direct copies of computer-mediated communications of online community members and reflective notes the researcher makes during the observation. The first kind of data is often plentiful, easy to obtain, and almost automatically transcribed. Therefore, the researcher's choices on what data to collect are particularly important, and should be guided by the research questions. The second kind of data, the reflective field notes, are a time-tested and highly recommended way to provide context to the data. The analysis of data collected using netnographic techniques does not differ from normal ethnography.

Analysis in ethnographic research is usually qualitative and based on a holistic view developed in intense contact in the field. Data is captured from the inside, in natural settings. Groundedness to local knowledge and long-term exposure to the field make it possible to study processes. In addition, this gives qualitative methods strong potential for testing hypotheses (Miles and Huberman 1994). Qualitative analysis is mostly carried out with words. Many interpretations of the data are possible, but some are more compelling (Miles and Huberman 1994). The end product of ethnographic research is a holistic, context-sensitive narrative of the everyday life of the social group. It is essentially two stories: one about the representation of results and the other about how that representation was constructed (Agar 2004). In order to carry out ethnographic research, it must be accepted that the researcher is a part of the story (Agar 2004). Field notes and observations are texts constructed by the researcher. They are influenced by his values and bias. Things also always happen in a context. The data speak more about actions people have taken rather than their behaviour in general. The critical assumption of ethnography and qualitative research is the researcher-as-instrument (Güney 2010; Miles and Huberman 1994). The researcher has a major role in data collection and analysis. Although the researcher carries a value system and all the bias that entails, he is also capable of critical reflection on his own influence on the interactions in the research setting. In terms of complex adaptive systems, the researcher is one of the agents making the system run; but, being only one of the many, he has only minor responsibility for the events that emerge (Agar 2004).

2.4 Research design

This study is conducted as a multiple case study according to the research design presented in Figure 2.1. First, the context of the study is clarified by reviewing relevant literature on innovation, innovation processes and crowdsourcing. Then a systematic literature review on collective intelligence is carried out. The results of the literature review guide the development of an initial theoretical framework. The theoretical framework is used to guide the ethnographic data collection on three cases. Collected data is first organized using qualitative data analysis software and then summarized by writing detailed case descriptions. These descriptions work as analytical tools, presenting the collected data in condensed format. The case descriptions are used as a basis for cross-case analysis. In cross-case analysis several theoretical lenses are used to compare the cases and draw conclusions about the role of collective intelligence in the investigated sites. During the cross-case analysis qualitative data is complemented with quantitative data collection and analysis, and a simulation model. The rest of this chapter is dedicated for presenting the general aspects of the data collection and analysis methods. For clarity, the details on specific data collection and analysis steps are discussed in each corresponding chapter separately.

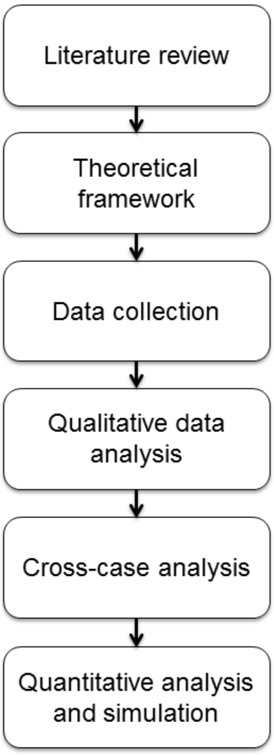


Figure 2.1: Overall research design.

2.5 Case study research and case selection

To investigate the role of collective intelligence in crowdsourcing, a multiple case study was conducted. Although propositions derived from existing literature are used to guide the research, the study is more focused on building an emerging theory than testing an existing one. Replication logic and cross-case comparisons are central when building theory from cases as each case serves as an experiment that contrasts and replicates the others. Emphasis is on the complex real-world context. As case studies remain close to the data, they can be both honest and objective (Eisenhardt 1989). Limited sample size is perhaps the most common criticism for case studies, but this is often misguided. Even a single case can make a powerful example (Siggelkow 2007). Multiple cases clarify whether a finding is idiosyncratic or can be consistently replicated (Eisenhardt and Graebner 2007).

Theoretical sampling is used to select cases, which means that cases are chosen for theoretical instead of statistical reasons (Eisenhardt 1989). Acquiring a representative sample is not the goal of case selection. In theoretical sampling, the cases are selected because they are suitable for illuminating the constructs of interest and their relationships. This is similar to laboratory experiments, which are not selected randomly from all possible experiments, but because of the high likelihood that the particular experiments chosen will provide theoretical insights (Eisenhardt and Graebner 2007). Valid reasons for selecting cases for a multiple case study include replication, extension of theory, contrary replication, and elimination of alternative explanations (Eisenhardt and Graebner 2007). The research approach for this study was inspired by 30-day challenges popularised in the movie *Super Size Me* (Spurlock 2004). Thirty days has been claimed to be a long enough period to develop habits (Babauta 2009), and challenges on various topics (e.g., Hudson 2015) are abundant. I investigated each case as a participant-observer for at least 30 days. The data was analysed mostly in qualitative fashion. The cases were selected based on their theoretical relevance and replication logic. The following criteria were used when selecting the cases:

1. Case sites should use crowdsourcing as a part of their innovation or product development process.
2. Case sites should be of high quality. The findings of the study should not be affected by poor implementation of crowdsourcing efforts.
3. Case sites should use similar approaches and processes for crowdsourcing to support the replication logic in analysis.
4. A priority was put on cases that do not require special skills from the participants.
5. Previous research has already suggested the possibility of collective intelligence on a particular site.

As the pool of potential sites is rather large (Crowdsourcing.org (2015) lists 2,885 examples of crowdsourcing sites as of 22 January 2015), the priority is put on the well-known sites. Using these criteria, three crowdsourcing cases were selected in the innovation and product development context: OpenIDEO, Quirky and Threadless. Each of these sites uses crowdsourcing as a part of their innovation or product development process. They are mentioned repeatedly as examples of crowdsourcing, are well-known, and crowdsourcing is part of their core business: The renowned design company IDEO hosts OpenIDEO. Quirky has managed to create rapid growth, investor interest and general hype about the company (e.g., Griffith 2013; Fenn 2012). Threadless is a classic example of successful crowdsourcing. The company was set up in 2000 and has been featured as a successful example of crowdsourcing numerous times in both academic and popular literature (e.g., Brabham 2008b; Brabham 2010; Hoyer et al. 2010; Malone et al. 2010; Pisano and Verganti 2008; Boudreau and Lakhani 2009). Many organisations that carry out crowdsourcing (e.g., Dell, Starbucks) do not rely on it for their survival, and thus their motivation on getting it right may be lower than in the selected cases. Each selected case uses at least superficially similar processes: the organisation publishes a challenge, participants submit content and evaluate it, and the organisation selects some of the submissions for further development. For instance, well-known open innovation site InnoCentive does not fulfil this criteria, because the submitted content is not visible to the crowd or evaluated by it. Participation in these sites does not require special skills. Although Threadless is focused on graphic design, it is possible to participate in evaluations and forum discussions in a meaningful way without graphic design skillsTable 1.1 Table 2.1 lists examples from previous research suggesting that these kinds of sites can manifest collective intelligence. Being analogous cases, the arguments for one should apply to the others. OpenIDEO was used as a pilot case to test and refine the data collection and analysis methods (Salminen 2013a). The researcher visited each site as a participant-observer on at least 30 days.

Table 2.1: Examples from previous research suggesting the existence of collective intelligence in the selected crowdsourcing sites

Case	Proposition and source	Source
OpenIDEO	'OpenIdeo.com and OpenPlanetIdeas.com are two similar collective intelligence sites which use crowdsourcing to solve some of the world's environmental and health problems and innovate new uses of technology'.	Paulini 2012
OpenIDEO	'If you know something that someone else doesn't, rather than cut them down as ignorant – take up the challenge of how you might thoughtfully help them up their knowledge. That way we build collective intelligence'.	OpenIDEO 2015
Quirky	'Our analysis [of Quirky] shows that a design process that includes collective intelligence shares processes of ideation and evaluation with individual and team design, and also includes a significant amount of social networking. Including collective intelligence in design can extend the typical design team to include potential users and amateur perspectives that direct the design to be more sensitive to users' needs and social issues, and can serve a marketing purpose'.	Paulini et al. 2011
Quirky	'Platforms like TopCoder.com, SecondLife.com, Quirky.com and GeniusCrowds.com are examples of online innovation using collective intelligence'.	Paulini 2012
Quirky, Threadless	Quirky and Threadless are classified as samples of collective intelligence systems with non-routine tasks and emergent output.	Yu et al. 2012
Threadless	'Collective intelligence (or CI) has recently emerged as a potential magnifier of design thinking. A surge of internet based social computing applications are achieving surprising results from people thinking collectively, without the aid or restrictions of formal organisation, supervision, or even payment in the conventional sense. Some of the best known applications, such as Threadless and Top Coder involve design activity'.	Murty et al. 2010
Threadless	'Crowdsourcing is an online, distributed problem-solving and production model already in use by businesses such as Threadless.com, iStockphoto.com, and InnoCentive.com. This model, which harnesses the collective intelligence of a crowd of Web users through an open-call format, has the potential for government and non-profit applications'.	Brabham 2010
Threadless	'Google. Wikipedia. Threadless. All are exemplars of collective intelligence in action'.	Malone et al. 2010
Threadless	Threadless is included as an example of using collective intelligence to generate and evaluate potential solutions.	Bonabeau 2009

2.6 Data collection

Ethnography generally uses three data sources: participant observation, interviews and documents. In netnography, the focus is mostly on documents copied from the web during participation and notes made by the researcher regarding his observations. Selecting which data to collect is an important analytical decision and already a part of data reduction for the analysis (Miles and Huberman 1994). As large amounts of data are available on the web even in small-scale forums, dealing with information overload is an important concern (Kozinets 2002). Yin (2008) lists three principles to be followed in data collection for case studies: using multiple sources of data, creating a case study database, and maintaining a chain of evidence. The data collection procedure used in this study is shown in Figure 2.2. Appendix A provides the case study protocol used to guide the data collection.

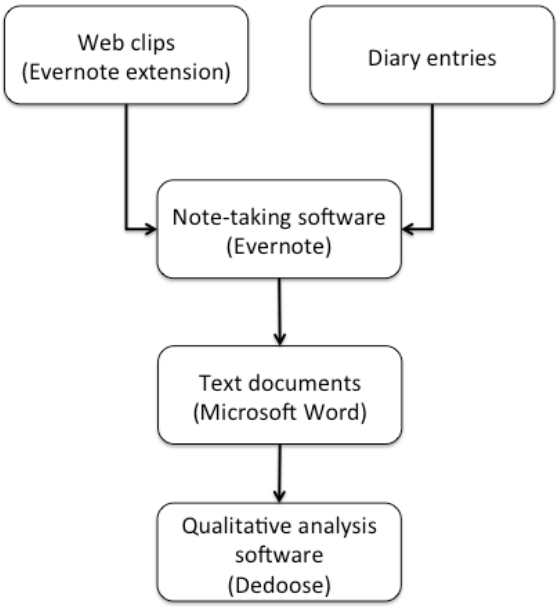


Figure 2.2: Qualitative data collection procedure used in this study

I used Notebook software (Evernote 2013a) and Evernote Web Clipper add-on (Evernote 2013b) on the Chrome web browser to collect interesting web pages that I visited during the participant observation. Ease of use allowed minimum distraction to participation due to data collection. The built-in functionality of the software helped to create an easily managed database. I ended up using two modes of data collection: usually I saved the pages on which I had spent some time or the pages I had shown interest in as a user. The resulting data are a sample of what users encounter. The sample is probably biased as I explored some less-used functionality, which I might not have done without the research interest. I documented my own observations in a diary, also stored in Evernote, where I noted all the major actions I took on the site, observations I made and feelings I had at the time. Diary entries varied from just a few lines to more than a page of text per field visit. Figure 2.3 depicts a sample note from the diary.

OpenIDEO, day 18

Checked the contribution: 81 views, 3 comments, 2 applause

Thinking about e-waste challenge

- Trying to figure out what exactly is the problem. Is it the design of products? Legal incentives? Consumers? Companies?
- Maybe the themes make figuring out the exact problem easier. Is that the purpose of themes? Breaking the problem to smaller sub-problems?
  - Make the business case is the most interesting theme for me. Checking the concepts under that theme.
  - There's none. Seems like people are not using the themes to organize the concepts. None of the theme filters provides any answers. Or maybe the functionality just doesn't work. My own concept has a theme, but it's still not showing up in corresponding theme. Shouldn't be a browser clitch, as it doesn't work on Safari either.
- Checked out the inspirations again. There the themes and missions work as filters. One of the missions was hosting an interview. Only one of the facilitators had done it. The results were presented in the blog.
- Figured out a possible concept by thinking about the problem and some of the inspirations I saw yesterday.
  - Tax on electronic products based on the time it takes to disassemble them for recycling. A modular design, where the toxic parts are easily removed from covering plastics would be fastest to disassmble. The same structure would probably be also easy to assemble, thus providing savings to producers. Tax could be calculated as a constant amount of money per second it takes to disassemble the product for recycling
  - Posting the concept this late will not give anyone opportunities to build on my concept. That's a shame. Should be posting concepts earlier.
- Uploaded a concept to E-Waste challenge: Disassembly time tax. Just in time, as there's only one day left before applause phase starts...

Figure 2.3: Example of a research note from the research diary.

I collected additional documents, such as toolkits for workshops and presentations of challenge results. I collected statistics on crowd evaluations from OpenIDEO and Threadless, but not from Quirky because the site did not provide open access to the data. I followed the data collection principles of Yin (2008). Web pages, diary entries, documents and evaluation statistics provide multiple data sources. I used Evernote to create and maintain a case study database and a chain of evidence, including dates of collection, web addresses and content. Table 2.2 presents a summary of cases, observation periods and data collected.

Table 2.2: Summary of the collected data. Data collection methods were refined during OpenIDEO case, and as a result the observation period was longer than in other cases.

Case	Observation period	Web clips	Diary entries	Statistics
OpenIDEO	51 days between 26 July 2012 and 24 December 2012	395	52	Views, comments and applause for three challenges before shortlist selection
Quirky	35 days between 2 Sep 2012 and 14 Sep 2012 and between 2 May 2013 and 28 May 2013	356	35	-
Threadless	30 days between 28 May 2013 and 30 June 2013	204	30	All scored designs in Threadless challenge between 24 July 2012 and 7 July 2013

2.7 Qualitative data analysis

Qualitative data analysis relies on three principles: data reduction as part of the analysis, use of data displays, and drawing and verifying conclusions based on these displays (Miles and Huberman 1994). As data reduction is a part of the analysis, the way in which it is done is an important analytical decision. There are many ways to reduce data. Anticipatory reduction limits the amount of data collected before the actual fieldwork through the selection of conceptual frameworks, cases, research questions and data collection approaches. During data collection, data is reduced by coding, categorisation, clustering and partitioning, and by writing summaries and memos (Miles and Huberman 1994). This form of analysis sharpens, sorts, focuses and organises data so that conclusions can be drawn. The overall structure of the qualitative data analysis procedures used in this study is presented in Figure 2.4.

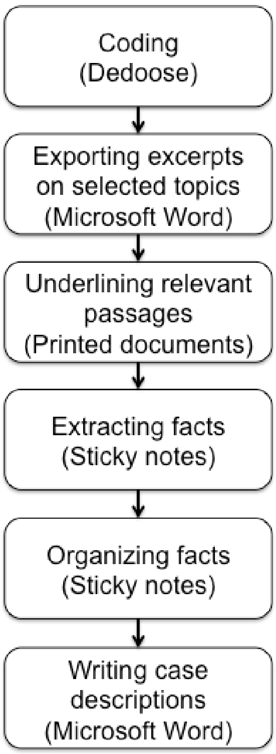


Figure 2.4: Qualitative data analysis procedures used in this study

Dedoose qualitative data analysis software (Dedoose 2013) was used to organise the data collected with Evernote. The notes were imported to analysis software as Microsoft Word documents. All data were coded using the code list presented in Appendix B. Codes are tags that assign meaning to chunks of data, such as words, phrases or paragraphs. They are used to organise data within a system of categorisation to facilitate retrieval of chunks of data relevant to particular research questions or themes. An initial list of codes should be created before the fieldwork begins, but the researcher should also maintain sufficient flexibility to refine the codes when they turn out to be inapplicable or ill-fitting to actual data (Miles and Huberman 1994). The initial code list was derived from the conceptual frameworks of collective intelligence, innovation processes and crowdsourcing. The codes evolved during the pilot study analysis: some were dropped, some added, and the use of some codes changed. Such variation in coding practices does not threaten the validity of results and is to be expected. One of the purposes of the pilot study was to develop a coding scheme and analysis procedures to be used in further cases. The coding was used to make retrieval of relevant data easier.

A good way to start the analysis of a case is to write an interim case summary (Miles and Huberman 1994). This is a provisional synthesis of what a researcher knows about the case, usually 10 to 25 pages in length. It provides the first coherent account of the case (Miles and Huberman 1994). After coding the data on Dedoose, the software was used to export selected data for further analysis using relevant codes. Excerpts both from web clippings and diary entries were included. The majority of data came from the web documents, expect for the code user experience, where the diary was a slightly more important source. The focus of the analysis was on tasks (activities), rules, feedback, and user experience (agents) because the theoretical framework suggests that these themes are important. Determining outputs of the system in different phases was straightforward, and as the websites functioned as the distributed memory, more detailed analysis of these themes was forgone. Inputs to the system come through the participants and consist of everything they have seen or experienced. They are therefore unknowable and were thus excluded from analysis. Human capabilities for interaction were outside the scope of this study because literature on psychology discusses them in much more detail than is possible here. Finally, emergence was not directly observable in the data but may or may not be revealed during the analysis and comparisons. The reduced datasets were read through and insights were collected on sticky notes, which were then clustered around emerging themes to reveal patterns in the data. Interim case summaries were written based on the patterns revealed by this analysis. Care was taken to use the same language, terms and phrases as used in raw data. The interim case summaries describe the operation of the site from the above-mentioned perspectives.

Extended text, even in the compressed format of a case summary, is cumbersome to use for analysis: the data tends to be dispersed, sequential, poorly organised and bulky. Therefore, valid analysis requires data displays that are focused enough to show the full dataset at once in a systematically organised format (Miles and Huberman 1994). This format makes it possible to draw conclusions. Miles and Huberman (1994) aptly underline the importance of displays: ‘You know what you display’. Displays can take the form of matrices, charts and networks. Good displays are designed to organise information so that it is immediately accessible, compact and allows the analyst to make careful comparisons, detect differences and note patterns and trends in the data. Drawing conclusions and verification consists of noting patterns, explanations, causal flows and propositions in the displays. At first, any conclusions should be tentative. An open and sceptical mindset is advisable. The meanings emerging from the data must be tested and confirmed, by seeking feedback from the stakeholders, for example (Kozinets 2002; Yin 2008), to ensure their validity and trustworthiness. The data displays used in this study are mostly based on the interim case description.

2.8 Cross-case analysis

As the concept of collective intelligence is still somewhat fuzzy, it is necessary to take into account several different interpretations of the phenomenon revealed in the literature review. Observed patterns gleaned from case descriptions are matched to theoretical patterns derived from different theoretical frameworks, progressing from general descriptions to more detailed examinations. This theory and method triangulation helps to increase the study's construct validity. The used theory lenses and methodological approaches include Collective intelligence genome, innovation as a search problem, properties of collective intelligence systems, wisdom of crowds, distributed cognition, and simulation modelling.

Collective intelligence genome is a framework for classifying collective intelligence systems developed by Malone et al. (2010). In this study the framework is used to characterize the investigated crowdsourcing sites in comparable terms, and then to define an analogical case to which the findings of the study should be generalizable. As discussed in the chapter 3.1, in the abstract level innovation can be viewed as a search problem in multi-dimensional space. This view lends itself to the comparison of search processes of individual inventors and collective intelligence systems formed by crowdsourcing sites and their participants. Schut (2010) has defined a set of properties that help identifying collective intelligence systems. This set of enabling and defining properties is used to identify the innovation process phases most likely to manifest collective intelligence on the investigated crowdsourcing sites. Wisdom of crowds refers to the improved accuracy of aggregated contributions from the crowd. This theoretical lens is used to evaluate the output of the collective intelligence systems in terms of accuracy of evaluations provided by the crowd. Quantitative data collection and analysis approach is used to compare the decisions made by the crowds and crowdsourcing organizations. The viewpoint of distributed cognition allows detailed examination of the crowds' interaction on the sites. Possible pathways of information are identified in the phases of the innovation processes deemed to have potential for collective intelligence in previous analyses. Finally, a simulation model is constructed and used to assess the effects the two different aggregation methods used on the case sites have on the accuracy of the crowd.

Table 2.3 lists the most relevant displays and the purposes for which they were created. Details of the methodological choices regarding the theoretical lenses are presented in corresponding chapters along with the descriptions of the results.

Table 2.3: Main data displays created during the qualitative analysis

Display	Purpose	Data sources	Notes
Innovation process	Description of the innovation processes	Case descriptions	
Collective intelligence genome	Identification of interesting phases for further analysis	Case descriptions	Malone et al. (2010)
Creation of innovations as a search process	Comparison of crowdsourced innovation processes to inventor's search process	Case descriptions	Maggitti et al. (2013)
Properties of collective intelligence systems	Comparison of criteria for collective intelligence systems and cases	Case descriptions	Schut 2010
Possible paths of information	Identifying potential paths information could take in the investigated crowdsourcing systems	Case descriptions	
Wisdom of crowds statistics and visualisations	Evaluation of wisdom of crowds effect	OpenIDEO and Threadless	Statistical analysis provided in Appendices C and D

2.9 Quantitative data analysis: Statistical learning

In addition to the qualitative research methods described above, this study relies also in quantitative methods in cross-case analysis. Quantitative data analysis entails collection and analysis of numerical data (Bryman and Bell 2007). More specifically the approach taken can be described as statistical learning. Statistical learning refers to an extensive set of tools and methods used for understanding data. James et al. (2013) divide these tools into two categories: supervised learning and unsupervised learning. In supervised learning, a statistical model is created to predict or estimate output values based on one or more inputs. In unsupervised learning, there is no output to predict but instead the goal is to learn about structure and patterns in data, for example, by looking for groups of similar inputs. The statistical tools used in this research fall mainly into the category of supervised learning. Technically, supervised learning models work by estimating a function, f, that connects the inputs to outputs. The function f represents the systematic information the inputs provide about the outputs, which can be used for making predictions and inferences about the data. In predictions, the interest is in the accuracy of the predictions made by the model, whereas what is going on inside the model is not very important. The model can be treated as a black box that converts the inputs to output predictions. In inference, understanding how the inputs affect the outputs is the main concern. The accuracy of the predictions made by the model may not be as important as the interest in the internal workings of the model: which inputs are associated with the output, what their relationships are, and how the relationships can be summarised.

Another way to categorise statistical tools is by grouping them to parametric and non-parametric methods. Parametric methods make an assumption about the functional form of the model and simplify the model-building problem to one of estimating the parameters of the function based on the training data. Although this is much easier than fitting an arbitrary function, the flexibility of the model naturally suffers (James et al. 2013). If, for instance, a linear relationship between variables is assumed, as in linear regression, the model is likely to fit the data poorly if the real relationship is more complex. Interpreting the meaning of parametric models is relatively easy. Non-parametric methods do not make assumptions about the functional form and can therefore fit even very complex relationships. Flexibility comes with costs, however. Non-parametric methods typically require much more data than parametric methods, and interpreting the models can be challenging. For instance, random forests make predictions of outputs by fitting several different decision trees to data and taking a vote on the individual trees. Although the method works well in making predictions based on complex data, it is not easy to deduce what is going on within the ensemble of models.

Selecting the best approach is a major challenge in data analysis (James et al. 2013). Statistical methods make trade-offs between the flexibility and interpretability of the models they produce. Statistical methods can make use of both quantitative and qualitative variables. Usually, the type of input variables is not important as most methods can deal with both types without problems. The type of output variable is important and affects the choice of analysis methods. Linear regression is suitable when the output is quantitative, such as salary or body weight, but performs poorly when the output is qualitative, such as gender or marital status.

In order to select a model that performs well, a way to measure its accuracy is needed. Existing measures all try to assess how well model predictions match the actual data. So-called mean squared error is one of the most common for quantitative settings. Model accuracy should always be assessed on unseen data, that is, data that was not used in training the model, as there is no guarantee that a model with smallest training set error will also perform well on a test set consisting of new data. The issue stems from a trade-off between variance and bias, the two competing properties of statistical learning methods (James et al. 2013). Both bias and variance decrease the accuracy of a model, but for different reasons. Variance measures how much the function f used to model the relationship between inputs and outputs would change if it was estimated using different training data. The more flexible the method, the more likely it is to have high variance. Linear regression assumes that the relationship between inputs and outputs is linear, which is often not the case in the real world (James et al. 2013).

In qualitative settings, the same general rules apply but the error rate is the most commonly used error measure. It measures the proportion of classification errors made by the model. However, error rate can sometimes be a trivial measure of accuracy. Where one of the classes (in binary output context) is rare in the data, a model can achieve a very low error rate by always predicting the more common class (James et al. 2013).

2.10 Quantitative data analysis process

Conducting a statistical analysis typically consists of the following tasks. After the data relevant to the research question has been obtained, it is usually necessary to clean and manipulate it before it is ready for analysis. Cleaning and manipulating data may include fixing errors in the data, converting measurements to different units, combining or splitting variables and changing the format of the data. This is an important task because poorly formatted data makes the statistical analysis difficult or may even prevent it completely where the software used does not support processing the data in its raw format. Data manipulation is time consuming, and is even said to take the majority of work effort in data analysis projects. A good structure for datasets is so-called tidy data, which makes further manipulation and processing of data easy. Wicham (2014) defines tidy data as a dataset where variables are stored in columns, observations are stored in rows and each file contains data on one type of experimental unit. Tidy data structure is used to store the statistical data used in this study.

When the data is cleaned and in proper format, the statistical analysis can begin. First, it is important to explore the dataset to gain an overall view on its properties. Awareness about the distribution of variables and the presence of possible outliers and rough relationships between the variables guide the selection of appropriate analysis methods and help avoid and diagnose possible problems during the analysis. Exploration consists of the investigation of the descriptive statistics of the dataset including the ranges, medians and means of variables. A visual exploration is made by plotting the data on suitable graphical displays. The aim of the data graphics is to display data accurately and clearly (Wainer 1984). It is a matter of substance, statistics and design (Tufte 2009). Strict rules for presenting data do not exist. The representation of numbers in graphical displays should be directly proportional to the numerical quantities they represent. Commonly used display types should be favoured and unnecessary decorations avoided. Labelling should be clear, full and consistent (Wainer 1984).

The actual statistical analysis consists of training one or more models, assessing their performance, diagnosing possible issues, and interpreting and verifying the results. A good practice is to split the available data randomly into training and test sets. The training set is used to fit the model while the test set is held back and used only in the end to assess the accuracy of the final model. Statistical significance or p-value is used to assess the existence of a systematic pattern in the data. It measures the probability of getting data equal or more extreme than the one at hand, if there was is no relationship between the variables of interest (Coe 2002). For instance, if the p-value of a statistical test is 0.1, it means that the there is a 10 % change chance for a random process to generate similar data than to the one used in the analysis. Traditionally, p-values smaller than .05 are considered to be statistically significant, but the practice has lately come under criticism. The probability of a research finding being true may depend on the power and bias of the study, the number of other studies looking at the same question, and the ratio of true relationships to the number of relationships being studied within a scientific field. As the results of statistical analyses are probabilistic, a number studies looking at the same question increases the possibility that one of them happens to be statistically significant by chance alone. It is possible to prove that most research findings reported in most scientific fields are actually false (Ioannidis 2005). The situation may be improved by conducting larger studies, by higher standards on research findings and by the use of a range of effect size measures in addition to reporting the significance levels of studies. Johnson (2013) suggests stricter significance levels of $p < 0.005$ for statistically significant results, with $p < 0.001$ for highly significant results as new acceptable significance levels. The goal is to avoid reporting spurious results as significant. However, any definition of acceptable significance levels is by nature arbitrary. It is partly the reader's responsibility to evaluate whether a particular statistical finding is believable. An issue with using significance levels for assessing relationships is that the measure confuses the size of the effect with sample size. In practice, almost any difference between two variables will appear statistically significant, given a large enough sample size. Therefore it is important to report additional measurements of statistical effects in addition to the significance levels of results (Coe 2002; Field et al. 2012). Effect size is a simple and easily understood method for quantifying the difference between two groups. Several alternative ways exist for calculating the effect size. Using at least one of these helps to assess whether a statistically significant finding is important and meaningful in practice.

Finally, the current statistical analysis software allows full reproducibility of research through sharing of raw data and analysis scripts. This functionality should be used whenever possible. The idea is that given the software packages, original datasets and computer code another researcher can run exactly the same analysis and arrive at exactly the same results (Peng 2009). Instead of trusting the condensed description of methods in research reports, the reader can investigate every step of the analysis in detail if he or she so wishes. The statistical analyses in this study were conducted using R programming language, which offers great support for reproducible research. This study strives to follow the comprehensible guidelines for the use of statistical methods offered by the APA Board of Scientific Affairs (Wilkinson et al. 1999). The statistical analyses conducted in this study are fully reproducible. Anonymised data and R scripts required for conducting the analyses have been published on GitHub. All data and intermediate results used during the qualitative analyses have been made available. These items include web clippings and diary entries exported from Evernote, coded data from Dedoose, excerpts related to rules, tasks, feedback and user experience, extracted facts organised with sticky notes, and case descriptions. Although the qualitative analysis steps are not easily reproduced, the data provided establish a continuous chain of evidence from raw observations to final conclusions.

2.11 Simulation modelling

As a final step of analyses carried out in this study a simple simulation model is constructed. Despite the usefulness of the static qualitative and quantitative methods described above, models using them can be less than dynamic. They omit feedback loops, delays and other nonlinearities typical of complex systems. Often the only practical way to test more dynamic models is simulation because the complexity of even simple models exceeds the human capacity to understand their implications. In addition, testing the models in the real world is slow and often inaccurate. The constraints and rigour required for building simulation models also make it more difficult to base the mental models on unfounded ideology or unconscious bias. Simulations, however, can omit important variables to keep the model tractable or because of the unavailability of exact data for estimating the model parameters. Broad model boundaries are much more important than level of detail to capture the proper dynamics of the system of interest. Simulations should never be built to model all the details of the real world system. Rather, the purpose of simulation should always be modelling a problem. The problem works as a guide and helps decide what aspects are important and what parts can be omitted. As a result, problem definition and characterisation is the most important step in simulation modelling. Reference modes, or sets of graphs and other data presenting the behaviour of the system over time are a useful tool at this stage. After identifying and characterising a problem, a provisional dynamic hypothesis should be developed to account for the problematic behaviour. Simulation modelling is usually used to search for endogenous explanations, that is, explanations that give rise to observed behaviour from within the system. The system is described on a conceptual level using various available system-mapping tools, and then formalised with simulation software. Finally, the model is tested. The simulation's behaviour is compared to that of the real system, including the extreme boundary conditions to ensure that the responses of the simulation are logical and coherent. The simulation model can then be used to generate and evaluate policy suggestions (Sterman 2000).

2.12 Research ethics

The unobtrusive nature of ethnography on the web, or netnography, makes this approach both attractive and controversial (Kozinets 2002). Should online forums be considered as a private or public space, and what constitutes informed consent in this context? In order to ensure ethical netnograpic research, the researcher should disclose his presence, affiliations and intentions to online community members. The confidentiality and anonymity of informants should be ensured, and permission should be obtained to use specific quotes and idiosyncratic stories in the research (Kozinets 2002). The ethical issues of this study are not very prominent as the focus of the research is on the processes, not so much on the behaviour of individuals. Direct quotes or personal stories attributable to individuals are not used in the report. I announced my identity and affiliations as a researcher on my profile page. As a result, I consider the research to follow ethical guidelines.

3 State of the art

This chapter provides a background to the study by reviewing relevant literature on innovation and innovation processes, crowdsourcing and collective intelligence. Literature on innovation and crowdsourcing provides the context of the study; the main focus is on collective intelligence. Accordingly, only the literature on collective intelligence is reviewed systematically, as described in below in 3.3. Furthermore, three recent systematic literature reviews on crowdsourcing were already available, reducing the need for replicating the work. The literature on innovation and crowdsourcing was selected for review based on an active following of the relevant literature streams during the research project. NAILS, an automated tool for literature analysis, was used to check that the most important works on the relevant fields had not been ignored. Keywords used in separate searches on the database included "collective intelligence", "crowdsourcing" and "innovation process". After examining a set of literature records downloaded from Web of Science Core Collection in tab-delimited format, NAILS produces a report displaying the most important authors, publications and keywords, and lists the references cited most often in the analysed papers (Knutas et al. 2015).

3.1 Innovation and innovation processes

Before rushing forward with the investigation, it is useful to spend some time charting the context and assumptions of the investigation. The review of the literature on innovation is necessary for understanding the rest of the thesis. Instead of trying to cover the whole field comprehensively, the aim is to establish a working definition of innovation and a representative description of innovation processes. The term innovation is used frequently in news, politics, academic literature and company strategies, but what does it actually mean? One way to go about clarifying a phenomenon is to construct a model of it. This approach is recommended as the first step of any investigation in a classic book on qualitative research, the *Handbook of Qualitative Data Analysis* by Miles and Huberman (1994). However, the best explanation of the usefulness of models is arguably the one given by Scott Page at the beginning of the *Model Thinking* course on Coursera (2012). According to Page, understanding models is useful for four reasons. First, models are everywhere, nowadays; it is nearly impossible to be an intelligent citizen of the world without understanding them. Secondly, models clarify thinking by forcing one to consciously select what is important about a phenomenon. The model also makes these selections explicit, making it easier for others to evaluate and criticise the thinking behind them. Making models may thus be risky for one's reputation, but at the same time it can be a form of honesty. Thirdly, models support the use and understanding of data. With big data being one of the buzzwords of our time, it is easy to find spurious correlations between unconnected phenomena (Ioannidis 2005). Having a model can help mitigate these risks. Finally, models support decision making, strategising and planning. A model can help predict the effects of different decision options and thus allow decision makers to make a more informed choice. A good model can condense large amounts information, as exemplified in a famous quote from physicist Richard Feynman:

> If, in some cataclysm, all scientific knowledge were to be destroyed, and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words? I believe it is the atomic hypothesis (or atomic fact, or whatever you wish to call it) that all things are made of atoms – little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another. In that one sentence you will see an enormous amount of information about the world, if just a little imagination and thinking are applied (Feynman et al. 1964).

This sentence could be called the simplest useful model of physics.

3.1.1 Definition of innovation

So what would be a similarly simple definition of innovation? The following is a good candidate:

Innovation = problem + implemented solution

It is assumed that the problem, the solution, or their combination is new. This definition is not original, as its spirit has been echoed widely in the literature. For instance, according to Ulrich (2011), 'design is conceiving and giving form to artefacts that solve problems'. Next, imagination and thought are applied to find out what information can be derived from such a simple definition. Problem and solution in this case should be understood in the broadest sense. Synonyms for problems include such concepts as need, latent need, requirement and job-to-be-done (Christensen et al. 2007). Similarly, the solution can be loosely understood as a product, service, idea or invention. This definition suggests that both a problem and a solution are needed for an innovation. If the problem is missing, the innovation is just an invention, a solution to a problem nobody really cares about. Conversely, the lack of a solution leaves the world with an unmet need. Casual observation suggests that people tend to pay much more attention to the solution than the problem (Murray 2010). Getting the problem or customer need wrong could even be one of the most common reasons for failures in startup companies (Blank and Dorf 2012). This definition reminds us that problems and solutions are perhaps equally important for innovation. The problem and solution should obviously fit each other, but they are also partially independent of each other. For instance, it is quite common to create new innovations by using old solutions to solve new problems (Murray 2010; Johnson 2010).

This definition of innovation, however, does not tell us everything. For instance, it does not tell us whether the problem or solution came first. Both cases are possible, and the history of innovation gives many examples of each. A large part of engineering science is about people developing novel solutions to known problems, such as travelling to the Moon, or landing a rocket in one piece on a small floating platform (SpaceX 2014). Xerox PARC used to stockpile many useful solutions but, partially due to its strategy at the time, the company was unable to identify which problems they could solve. Eventually the solutions were matched with appropriate problems, but not by the company that developed the solutions in the first place (Chesbrough 2003). In addition, the definition of innovation described above does not tell us how innovations are created; it merely identifies the elements to be found in all innovations. Nor does the definition suggest whether the person (or other entity) needs to be conscious that innovation consists of a problem and a solution.

3.1.2 Different theories and viewpoints on innovation

The definition of innovation as problem plus implemented solution is simple, and easy to remember and understand. It also captures a significant part of what is essential about innovations. As an organising principle, it helps put innovation research in context and can thus help in understanding various theories on innovation. Next, we compare this definition with different theories and viewpoints on innovation, including lead user theory, open innovation, fitness landscape models, and anecdotes about product development in different companies. The idea is to give a brief overview of the innovation theories relevant to this thesis and to illustrate the applicability of the innovation definition in various contexts and frameworks.

Lead user theory argues that many valuable innovations are created by so called lead users. According to von Hippel, who developed the theory, lead users are those whose current needs or problems will become general and widely present in the marketplace in the future (Hippel 1986). Put another way:

> They are at the leading edge of an important market trend(s), and so are currently *experiencing needs* that will later be experienced by many users in that market. They anticipate relatively high benefits from *obtaining a solution* to their needs, and so may innovate (Hippel 2005, emphasis added).

A classic example of lead users in action is the story of the birth of mountain bikes. In the 1970s, biking enthusiasts discovered that it was enjoyable to ride bikes down steep hills. The only problem was that the bikes of the time tended to break down under such demanding conditions. As a result, users modified their bikes by strengthening the frames and adding parts from motocross bikes. In this way, they created the first mountain bikes. Only later did bike manufacturers discover that there was a market for sturdy bikes and develop commercial models. Currently, mountain bikes comprise a market worth several billion euros. A solution that worked for riding down mountains, a relatively niche market, was also found to be useful on rough cobblestone streets, a much wider market (Hippel 2005).

Open innovation is a concept popularised by Henry Chesbrough. It is defined as:

> … a paradigm that 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. Open Innovation combines internal and external ideas into architectures and systems whose requirements are defined by a business model (Chesbrough 2003).

This definition is mostly concerned with how the innovation happens rather than what an innovation is. What open innovation means in practice is illustrated with an early example of the use of open innovation in an organisation. Huston and Sakkab (2006) describe how Procter and Gamble adopted principles of open innovation in the company's approach to product development:

> We created a technology brief that defined the *problems we needed to solve*, and we circulated it throughout our global networks … to discover if anyone in the world had a *ready-made solution'* (Huston and Sakkab 2006, emphasis added).

Another example of the phenomenon is InnoCentive, a platform for open innovation. According to the company's own description:

> InnoCentive is the global leader in crowdsourcing *innovation problems* to the world's smartest people who compete to provide *ideas and solutions* to important business, social, policy, scientific, and technical challenges (InnoCentive 2013, emphasis added).

In open innovation, it is often the case that the system owner defines the problem and open participation provides the solutions.

In more abstract terms the creation of innovations can be framed as a recombinant search over the space of possible solutions (Fleming and Sorensen 2001). The space of possibilities is naturally enormous; accordingly, it is not possible to try out every solution. Instead, some other approach must be used. While search algorithms on computers often rely on approaches such as local hill-climbing (Russell and Norvig 1995), this is not how humans typically create innovations. Rather, the individual inventors[1] rely on more flexible search heuristics (Maggitti et al. 2013). Under the paradigm of innovation as search, it is assumed that searching for solutions locally (i.e., focusing on the neighbourhood of solutions already familiar to the searching agent) is easy compared to a distant search (searching for solutions not related to the current knowledge of the searching agent).

Although originally developed to study evolutionary biology, the metaphor of search may also be applicable in the context of innovations. The generation of scientific

[1] Inventor: a person that creates innovations.

breakthroughs has been described as recombinant search (Schilling and Green 2011). Hong and Page (2004) modelled problem solving as a search problem. The search metaphor has been used to model organisational development (Levinthal 1997) and the creation of technological improvements (Kauffman et al. 2000). The results from analysing patent databases support the applicability of the search metaphor to the process of invention (Fleming and Sørenson 2001). In crowdsourcing, the search for high quality contributions from a large mass of ideas has become a bottleneck in the innovation process (Alexy et al. 2011).

Maggitti et al. (2013) complement these high-level analyses by investigating how individual inventors actually perform the search process. According to their model, the inventors rely on five routines during their search process: stimulus, net casting, categorisation, linking, and discovering. Stimulus refers to the identification of a problem or opportunity: the inventor realises that there is a need that can be satisfied. Net casting is about gathering information both within and outside the domain of interest. Inventors use many sources of information and dynamically switch between focusing on details and the big picture. Information gathering needs high levels of motivation. The information and alternatives found during net casting are filtered in the categorising stage. The inventor integrates the new knowledge to an existing mental schema or knowledge structure. Often, the new information forces the inventor to update the mental categories, which may lead to linking. In the linking stage, the inventor integrates knowledge from seemingly different disciplines to come up with a new idea or hypothesis. Finally, in the discovery stage, the inventor tests the validity of the idea. These phases are summarised in Table 3.1. The search process is not linear, but involves iteration back and forth between the stages. System complexity, the search context and the expertise of the inventor all have an influence on the search process.

Table 3.1: Inventor's search process stages (Maggitti et al. 2012)

Process stage	**Definition**
Stimulus	Identification of problem or opportunity
Net casting	Gathering of information both within and outside the domain of interest
Categorising	Filtering of information revealed in the casting stage
Linking	Integrating ideas from seemingly different disciplines to arrive at discovery
Discovery	Testing and validating the idea

3.1.3 **Innovation process**

Now that we have figured out what innovation is, we can take a look at how innovations are created by a brief review of the literature on innovation processes. Although complex and iterative by nature, the design and creation of innovations often follow the same process, albeit loosely. Various frameworks have been developed to model innovation processes, which are usually described as multi-stage processes with feedback loops between stages. A number of these descriptions are compared in Table 3.2. Figure 3.1 illustrates the typical phases of the innovation process.

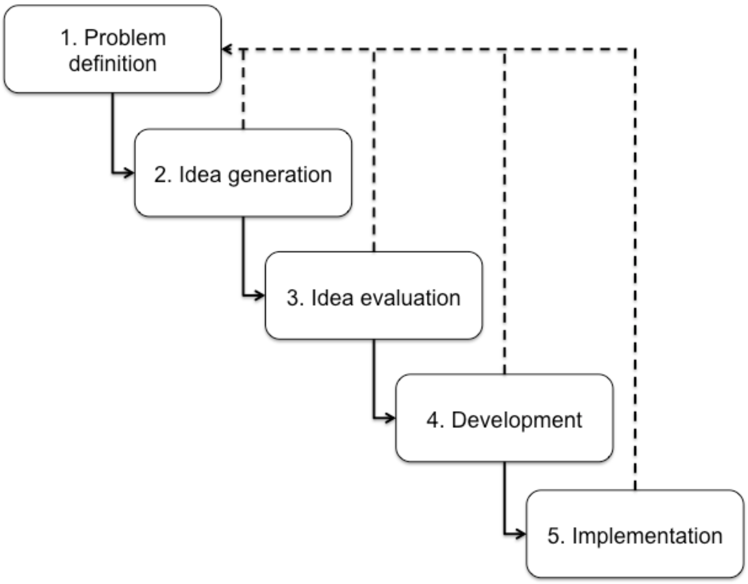


Figure 3.1: typical description of an innovation process. Note the close relationship to the definition of innovation: Phase 1 relates to a problem; Phases 2–4 relate to the solution; and Phase 5 relates to implementation.

The innovation process usually begins with an implicit or explicit problem definition. Some models place a strong emphasis on this phase (Kumar 2009), whereas, in others, the process starts with idea generation (Desouza et al. 2009; Cooper 1990). In these cases, problem definition is only implicit. Problem definition relates to learning about the environment, technologies (Veryzer 1998) and user needs (McFadzean et al. 2005). Learning can be a passive scanning of the environment for relevant signals (Tidd et al. 2005) or active research on the needs, hopes and issues of users (IDEO 2009). Most models see idea generation as a separate phase of the innovation process. New ideas are created to form a basis for further development. Again, this part of the process can be either explicit or implicit, as exemplified by the use of brainstorming (IDEO 2009) and the emergence of a vision about the possibilities of new technology (Veryzer 1998). The relative location of idea generation in the innovation process varies throughout different models. The innovation process may start with the generation of an idea (Desouza et al. 2009; McFadzean et al. 2005), just after (Cooper 2002) or long before it (Kumar 2009; Veryzer 1998; IDEO 2009). The number of initial ideas is often very large; in idea generation, quantity is considered to be more important than quality (IDEO 2009).

After idea generation the numbers of ideas in the process are reduced through evaluation and selection. Depending on the context, the focus can be on technological aspects

(Veryzer 1998), human aspects (IDEO 2009) or economic aspects (Cooper 1990). The most promising ideas are refined and combined. During this process, user requirements and technological features are defined more rigorously. Preliminary designs and even some early prototypes can be developed to clarify ideas (Veryzer 1998). The end result of this phase is the concept, which will be turned into reality in the next steps of the process. The development phase is executed based on the concept (Tidd et al. 2005). Much experimentation takes place as more and more comprehensive prototypes are built and tested for their technological functionality and user acceptance (Desouza et al. 2009; Veryzer 1998; McFadzean et al. 2005). Expected and unexpected problem solving loops are characteristic: most of the costs are generated in this phase (Tidd et al. 2005). The viability of the project is tested from the perspectives of product, production process, customer acceptance and finance (Cooper et al. 2002). Manufacturing processes are designed and marketing becomes increasingly involved in the project (Cooper 1990). In the implementation phase, the innovation is more or less ready; technological issues have been solved and the current prototype works as required (Veryzer 1998; McFadzean et al. 2005). The business plan is now implemented, manufacturing is ramped up and full-scale marketing to customers begins (Cooper 1990; McFadzean et al. 2005). The innovation is launched on the market or taken into everyday use (Tidd et al. 2005; Shaw et al. 2005). Although different models use different terms to describe the phases, emphasise different aspects of the process or have divided the process in varying ways, there is substantial consensus in the literature about the structure of the innovation process. In short, innovation processes are about identifying a problem, searching for a solution and putting the solution into practice. It is assumed that the elements of the process described above can also be found on crowdsourcing sites that focus on the creation of innovations.

Table 3.2: Comparison of innovation process descriptions found in the literature

<table>
<tr><td></td><td>Problem</td><td colspan="6">Solution</td><td colspan="4">Problem-solution fit</td></tr>
<tr><td>Model</td><td>Problem definition</td><td>Idea generation</td><td colspan="2">Idea evaluation</td><td colspan="3">Development</td><td colspan="4">Implementation</td></tr>
<tr><td>Innovation process (Tidd et al. 2005)</td><td>Learn</td><td>Search</td><td colspan="2">Select</td><td colspan="7">Implement</td></tr>
<tr><td>Stage-gate process (Cooper et al. 2002)</td><td></td><td>Discovery</td><td>Scoping</td><td>Build business case</td><td colspan="2">Development</td><td>Testing and validation</td><td colspan="2">Launch</td><td colspan="2">Post-launch review</td></tr>
<tr><td>Summary of innovation processes (McFazdean et al. 2005)</td><td></td><td>Idea generation</td><td colspan="5">Problem solving</td><td colspan="4">Implementation and diffusion</td></tr>
<tr><td>Organisational innovation process (Desouza et al. 2009)</td><td></td><td>Idea generation and mobilisation</td><td colspan="2">Advocacy and screening</td><td colspan="2">Experimentation</td><td colspan="2">Commercialisation</td><td colspan="3">Diffusion and implementation</td></tr>
<tr><td>4th generation innovation process (Rothwell 1994)</td><td>The needs of society and state of the art technology</td><td>Idea generation</td><td colspan="4">Research, design and development</td><td>Prototype production</td><td colspan="3">Manufacturing</td><td>Marketing and sales</td></tr>
<tr><td>Human-centred design process (IDEO 2009)</td><td>Hear</td><td colspan="3">Create</td><td colspan="7">Deliver</td></tr>
<tr><td>Radical innovation process (Veryzer 1998)</td><td>Dynamic drifting</td><td>Convergence</td><td colspan="2">Formulation, preliminary design and evaluation preparation</td><td>Formative prototype</td><td colspan="2">Testing and design modification</td><td colspan="4">Prototype and commercialisation</td></tr>
<tr><td>Technological innovation process (Cantisani 2006)</td><td>Marketing</td><td>Conception</td><td colspan="2">Applied research</td><td>Experimental development</td><td colspan="2">Engineering</td><td colspan="4">Marketing</td></tr>
<tr><td>Design process (Ulrich 2011)</td><td>Sense gap, define problem</td><td>Explore alternatives</td><td colspan="2">Select plan</td><td colspan="7"></td></tr>
</table>

3.2 Crowdsourcing

Crowdsourcing refers to outsourcing tasks traditionally performed within an organisation to an undefined crowd, usually through an open call posted to the Internet (Howe 2008). Crowdsourcing as a concept became widely known after Jeff Howe's article 'The Rise of Crowdsourcing' was published in *Wired Magazine* in 2006. According to the article, crowdsourcing is similar to outsourcing, but instead of using a selected and predefined provider, the task in hand is outsourced to undefined crowds (Howe 2006). The article describes several recent examples of companies that had created marketplaces for latent human talent. These markets allowed anyone with the right skills to participate or sell those skills. The importance of the phenomenon was increased because these companies were often able to offer their crowdsourced services at lower prices or more efficiently than their competitors. The emergence of crowdsourcing is largely attributed to technological advances, especially in the digital domain, which has made advanced tools (from product design software to digital cameras) available and affordable to amateurs and hobbyists in addition to professionals. Crowdsourcing applications are found in very different fields, ranging from small routine tasks to photography and design services, science, public policy making, and innovations (Howe 2006, Doan et al. 2011, Aitamurto and Landemore 2013). Early examples include Amazon Mechanical Turk, iStockphoto, FoldIt, InnoCentive, and Iceland's experiment in crowdsourcing its new constitution. Amazon Mechanical Turk is one of the early examples of crowdsourcing. Initially developed as an internal tool, the service offers a marketplace for so called human intelligence tasks (HITs), small and often repeatable tasks that are difficult to automate. At the time of writing, HITs available at the site include extracting purchased items from shopping receipts, categorising products and tagging images (Amazon Mechanical Turk 2015). Workers typically get paid anything from a few cents to tens of cents per completed task.

iStockphoto is a stock photography company that created a marketplace for amateur photographers. Despite its humble starting point, the company gained market share from expensive professional stock photograph services. Whereas a traditionally produced stock photograph may cost hundreds of euros, iStockphoto offered photographs with prices starting from a single dollar. The quality may not have been on a par with the more expensive products, but for many it was sufficient: the pricing was irresistible (Brabham 2008a). The growth of the company caught the eye of industry members. In the end, Getty Images bought iStockphoto for about 50 million dollars.

FoldIt is an online game where the players are tasked with folding proteins to find the lowest-energy configuration. At the same time, the thousands of players help scientists explore the structure of important molecules. The approach demonstrated its potential when the gamers managed to identify the structure of a retrovirus enzyme that had baffled researchers for years (Khatib et al. 2011). The finding could help in the development of an AIDS vaccine.

InnoCentive is an open innovation intermediary started by Eli Lilly and Company. It helps companies solve problems they have been unable to solve with internal research and development. With assistance from InnoCentive, companies formulate their problems in general terms to make them understandable for as wide an audience as possible and then post them online for anyone to see. Participants propose solutions and the searcher company selects the winners, who receive a monetary prize in exchange for intellectual rights to the solution. The prizes can amount to tens of thousands of dollars.

Between 2010 and 2013, Iceland used crowdsourcing as a part of its constitutional reform efforts. This use of crowdsourcing was an historical milestone because it was the first time that the direct participation of the people was used in drafting a nation’s foundational document (Landemore 2014). Although the reform eventually failed, as the Icelandic parliament did not pass the bill, the case still demonstrates the scale of expectation that has at times been placed on crowdsourcing.

3.2.1 Definition of crowdsourcing

Since the original article, both practical and academic interest in crowdsourcing has risen considerably. Figure 3.2 shows the number of articles on the topic published per year on Web of Knowledge. The crowdsourcing industry site, Crowdsourcing.org, lists 2,885 examples of crowdsourcing sites as of 22 January 2015 (Crowdsourcing.org 2015). Although defining crowdsourcing explicitly and exactly is challenging, one approach is to frame crowdsourcing as a general-purpose problem-solving method: a crowdsourcing system enlists a crowd of humans to help solve a problem defined by the system owners (Doan et al. 2011). Estellés-Arolas and Gonzáles-Ladrón-de-Guevara (2012) give a more detailed definition:

> Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken.

Both of these definitions is usable for the purposes of this thesis. The investigated cases fit both definitions of crowdsourcing.

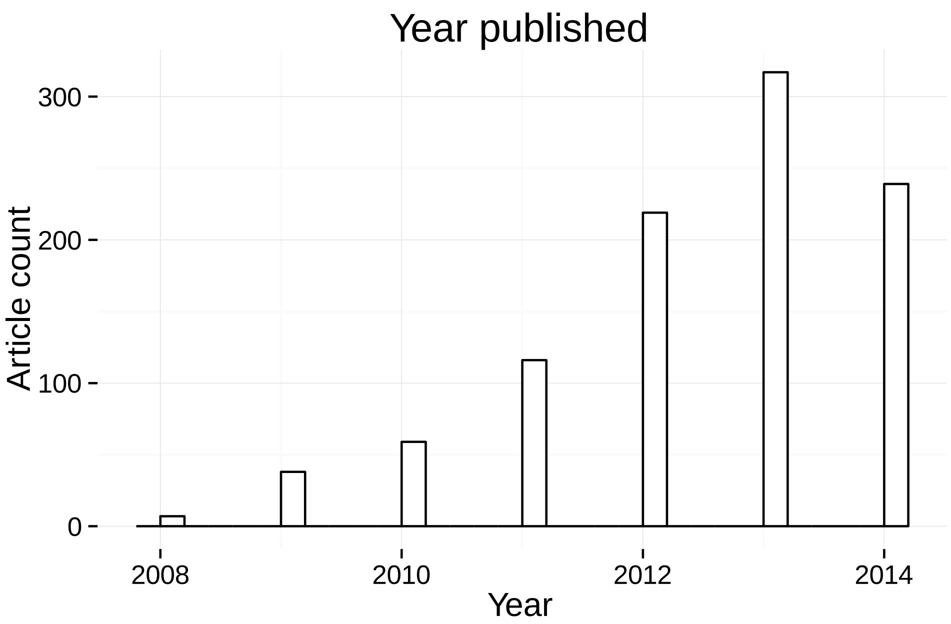


Figure 3.2: Number of published articles on crowdsourcing at Web of Science core collection (Web of Science 2014)

3.2.2 Promises and perils of crowdsourcing

Although crowdsourcing has been claimed to have the potential to bring about major change in society (e.g., Howe 2006; 2008; Malone et al. 2011), only preliminary investigations and speculation on the actual impact are available to date. According to Howe (2008), the rise of crowdsourcing has its roots in recent technological developments. Advances in digital tools and communication technologies are particularly significant. Both the means of production and distribution have become available to an ever-wider population. These developments may change the way work is organised. Malone et al. (2011) suggest that the work of the future may become more atomised, with more specialised workers taking care of smaller tasks. The driving force behind this development would be the economics of division of labour, which has been an important factor behind productivity gains since the industrial revolution. If crowdsourcing can – thanks to skilled amateurs and advanced communication technologies – offer solutions with similar quality to traditional work methods, but at a fraction of the cost, current organisational forms may need to be reshaped. New managerial skills would be needed to break down the work to smaller tasks, to assign the tasks to suitable workers, for quality control and to integrate the individual contributions into a coherent solution. At best, future organisations could adapt their capacity for routine tasks and tap the knowledge of experts for more creative endeavours (Malone et al. 2011).

The IT company, Apple, demonstrates the possibilities of crowdsourcing from the organisational perspective. According to Bergvall-Kåreborn and Howcrowft (2013), Apple's ability to rely on third party software developers is pivotal to value creation on its platform. Retaining highly skilled software developers on the payroll is expensive. Since 2008, the Apple platform has been open to software applications developed by outsiders. By crowdsourcing application creation, the company has been able to avoid the direct costs and associated risks of development, increasing its proportion of captured value in the process. On the other hand, most individual developers face intense competition and often find it difficult to achieve significant financial returns. The successful use of crowdsourcing in some companies may create complications in the other firms within the industry. The original *Wired Magazine* article (Howe 2006), already notes the tension between traditionally employed stock photographers and iStockphoto. Since then, discussion on the challenges to professionals offered by emerging communities of DIY designers has continued (Massanari 2012). At a minimum, crowdsourcing can affect professional identify as the relationship between consumers and producers changes (Aitamurto 2011).

For the workers, crowdsourcing promises possibilities for social mobility by offering new sources of income and by reducing traditional job market boundaries. Work flexibility may also be increased, as crowdsourcing allows workers to choose their own working hours and tasks (Malone et al. 2011). However, there are no guarantees of this positive future actually emerging. Payments for crowd work vary considerably. At the lower end of the scale, the picture looks bleak. Reportedly, working on micro tasks may pay as little as $2 per hour with no worker benefits are included (Kittur et al. 2013). In the worst case, crowdsourcing could displace current workers and substitute them with digital sweat-shops, where continuous, menial tasks must be performed for mere survival. Even workers with advanced skills face the risk of salary inflation. The common contest structure of many crowdsourcing sites removes payment guarantees (Malone et al. 2011).

As with any tool, crowdsourcing can be used for dubious purposes. Al-Qaeda is known to have used crowdsourcing to promote terrorism (Reid 2013). A robust ecosystem for circumventing CAPTCHAs (Completely Automated Public Turing Tests to Tell Computers and Humans Apart) already exists. These ubiquitous defence mechanisms are designed to protect web resources from illegitimate automated use (Motoyama et al. 2010). CAPTCHAs are small tests that are designed to be easily solved by humans but are very difficult for computers. They are commonly visual challenges, such as typing a sequence of distorted letters. When a spam bot is trying to sign up for a free email account but is blocked by a CAPTCHA, it can forward the puzzle to crowdsourced human workers and get the correct solution back before the sign-up page times out. The workers usually come from low-income countries and the payments are accordingly very low (Motoyama et al. 2010). This example demonstrates several potential crowdsourcing perils: misuse of web resources is facilitated by exploitation of the workforce. Kittur et al. (2013) suggest the development of an ethical framework in order to keep the conditions of crowd work acceptable for future generations.

Crowdsourcing is an important topic as it has the potential to bring about major change in our societies. However, there is no certainty about how these changes may take effect. On the one hand, crowdsourcing may ruin traditional industries, reduce salaries, lead to a deterioration in the working conditions of those who still have jobs, and facilitate illegitimate activities. On the other hand, crowdsourcing promises improved effectiveness and efficiency, both in routine and creative tasks, and simultaneously supports social mobility and economic benefits for the workforce. The situation is similar to current debates about the risks and benefits of rapidly improving artificial intelligence (AI) applications. Although we are constantly finding new areas where AI can offer significant improvements, the introduction of AI threatens up to 40% of jobs within the next few decades (Frey and Osborne 2013). It is important to increase our knowledge on crowdsourcing in order to make informed decisions on these issues.

3.2.3 Crowdsourcing as a research field

More detailed analyses on crowdsourcing literature have been conducted recently including work by Doan et al. (2011), Gabelloni et al. (2013) and Zhao and Zhu (2012). These studies are helpful in organising the research streams within the field. Research streams on crowdsourcing appear to be divided into conceptualisation, system and application focus (Zhao and Zhu 2012). Conceptualisation focus is about exploring what crowdsourcing is, how crowdsourcing works, and how it differs from related concepts, such as open innovation, outsourcing and open source. Conceptualisation is still a work in progress. Within system focus, crowdsourcing is seen as a set of interacting or interdependent components and the structure they form. Doan et al. (2011) provide a good overview on crowdsourcing systems and show that crowdsourcing can be applied to a wide variety of different fields and problems. Despite the differences in application domains, crowdsourcing systems typically consist of three categories of components: 1) the organisation that initiates the process and expects to benefit from the work; 2) the crowd of individuals who take care of the crowdsourced tasks; and 3) intermediating platforms (usually web-based) that connect the organisation to the crowd (Zhao and Zhu 2012). The platform also helps in clarifying the rules of the crowdsourcing effort to participants, managing the tasks and providing feedback to the crowd on their performance. Finally, the application focus explores specific crowdsourcing applications in different situations and for different purposes.

This thesis has both a conceptualisation and system focus: crowdsourcing is seen as a system, and the research is focused on how the crowdsourcing system works. Crowdsourcing systems can be understood in terms of the following fundamental dimensions: 1) staffing, or who is doing the task; 2) motivation and incentives, or why they are doing it; 3) goals, or what is being done; and 4) process, or how it is being done (Zhao and Zhu 2012; Malone et al. 2010). These dimensions are used to organise the previous research on crowdsourcing relevant to this thesis.

3.2.4 Motivation to participate in crowdsourcing

Why do people participate in crowdsourcing, especially when there are not always economic benefits in doing so? It turns out that the motivational factors are already quite well known. At the most fundamental level, human beings are probably motivated by the need to belong, or a strong desire for interpersonal attachment. People form social bonds readily in most conditions and resist the dissolution of already existing connections (Baumeister and Leary 1995). While existing evidence strongly supports the belongingness hypothesis, it is necessary to look at motivation in more detail to better understand participation in crowdsourcing.

Motivation can be either intrinsic or extrinsic. Extrinsic motivation refers to situations where an activity is performed in the expectation of an external reward. Activities that provide their own inherent reward are intrinsically motivated. People do these activities because they are satisfying in their own right. Research interest in both extrinsic and intrinsic motivation and their interactions has been keen over the last decades: a frequently cited meta-analysis has reviewed over a hundred studies investigating the effects of extrinsic rewards on intrinsic motivation (Deci et al. 1999). The overall conclusion is that external rewards tend to decrease intrinsic motivation towards an activity. Despite the common use of external rewards, especially in business, recent research has found that high and performance-contingent rewards can actually decrease performance (Ariely et al. 2011). Validated inventories are also available to assess the individual’s intrinsic and extrinsic motivational orientations (Amabile et al. 1994).

Although findings on human motivation should be widely applicable, the rise of crowdsourcing has created a stream of research on motivation in this specific context. For instance, it is known that a user’s decision to participate in a particular task depends on the number of other participants, the level of uncertainty in the task and the size of the reward. Fewer competitors, more uncertainty and higher rewards tend to increase participation (Boudreau et al. 2011). Several publications offer instructions on how to use rewards in crowdsourcing. Wagner (2011) suggests the rewards are particularly useful when it is necessary to attract unknown participants to solve a problem at hand, and when there is value in attracting multiple solutions. The problem to be solved and criteria for success should be clearly defined, and the agreed upon rules should be maintained throughout the crowdsourcing challenge (Wagner 2011). In this way, diverse risk takers and investors can be engaged and their trust towards the organisation maintained. More detailed guidelines (Finnerty et al. 2013) and a review of incentive mechanisms (Scekic et al. 2013) are also available. These articles provide a good starting point for designing incentive schemes for a crowdsourcing site.

In addition to extrinsic motivation induced by opportunities to make money or gain peer recognition, several intrinsic motivational elements in crowdsourcing settings have also been identified. A survey on participants at iStockphoto revealed that the sheer fun of the activity itself and the opportunities to develop skills are important motivational factors (Brabham 2008a). These factors may vary from site to site, however. A series of interviews with Threadless community members identified largely similar reasons for participation. In addition, many of the interviewees mentioned love of the community and addiction to the site as a major reason for continued participation (Brabham 2010). Overall, it has been suggested that some kind of balance between intrinsic and extrinsic motivation should be aimed at in crowdsourcing, intrinsic motivation perhaps being more important (Finnerty et al. 2013). We already know quite a lot about what motivates people to participate in crowdsourcing. As research on human motivation both in general and specifically in the crowdsourcing context is thriving, there is no need to investigate motivational aspects of crowdsourcing further in this thesis.

3.2.5 Crowdsourcing and innovation process

Overall current research indicates that crowdsourcing has the potential to contribute to the creation of innovations, as long as it is applied in the right circumstances (Aitamurto et al. 2011). According to the open innovation paradigm, organisations should use both internal and external sources of ideas and paths to market in their pursuit of innovation (Chesbrough 2003). Excluding closed partnerships, such as consortiums and subcontracting, there are two basic approaches on how companies can manage open innovation: communities and contests (Pisano and Verganti 2008; Boudreau and Lakhani 2009). In contests (or competitive markets), the innovating organisation publishes the problem and anyone can propose solutions. The organisation then selects the best. Governance tends to be formal. The external innovators are mainly extrinsically motivated and have mostly competitive relationships with each other. Direct contracts and transfers of intellectual property make value capture possible (Boudreau and Lakhani 2009). The communities are collaborative networks, where anyone can describe problems, propose solutions and decide which solutions to use. Compared to contests, communities are managed less formally. Participants tend to have more intrinsic motivations and more cooperative relationships with each other. Value capture from communities is also more difficult and may be possible only indirectly (Boudreau and Lakhani 2009). Open source software development is a typical example of the use of communities. In crowdsourcing innovations, the contest format is more common. Some recent examples include InnoCentive, a site where companies can post difficult problems for anyone to solve (Jeppesen and Lakhani 2010), Dell IdeaStorm, a website used to collect ideas for computer company Dell (Di Gangi and Wasko 2010), and My Starbucks Idea, where ideas are collected to improve the services and products of the Starbucks coffee shop chain (Sullivan 2010). The distinction between the two modes is not straightforward. Despite having features resembling a competition, such as the selection of winners, OpenIDEO describes itself as a collaboration (OpenIDEO 2012b). Similarly, although Pisano and Verganti (2008) classify Threadless as a contest, Boudreay and Lakhani (2009) emphasise the site’s community oriented aspects. Next, we will look in detail at what is already known about crowdsourcing tasks related to the creation of innovations using the contest format.

3.2.6 Crowdsourcing as a search algorithm

Afuah and Tucci (2012) propose an explanation for the effectiveness of crowdsourcing. They claim that under certain conditions crowdsourcing can improve problem-solving performance by transforming a distant search into a local search. Consider a cognitively limited problem solver. His or her position in the landscape reflects the information and knowledge the solver has access to. Instead of understanding the whole knowledge landscape, this solver is only familiar with the neighbourhood of his or her current position. Searching for solutions near the current position is relatively easy; this is called local search. Even a cognitively limited problem solver can perform well in a local search. On the contrary, extending the search outside the known neighbourhood in the knowledge landscape is difficult; this is called distant search.

Crowdsourcing can improve problem-solving performance in cases where distant knowledge is needed (Afuah and Tucci 2012). Instead of performing a distant search, the problem solver may crowdsource the task to a large number of participants, all of whom perform only local searches. If the number and diversity of participants are high enough, the chances are that someone is already in the neighbourhood of the solution, and can find the solution by a local search. Thus, distant search is transformed into local search and the performance of problem solving is improved.

3.2.7 Innovation contests

Following earlier definitions (Walcher 2007; Bullinger and Möslein 2010), Adamczyk et al. (2012) define innovation contests as 'IT-based and time-limited competitions arranged by an organization or individual calling on the general public or a specific target group to make use of their expertise, skills or creativity in order to submit a solution for a particular task previously defined by the organizer who strives for an innovative solution'. Basically contests connect two sets of actors: the sponsors and organisers of the contest, and the participants. Until recently the sponsor and organiser were the same, but nowadays the roles can be separated. A sponsor provides the prizes and defines the problem with the help of the organiser, whose main responsibility is the management and promotion of the contest itself. The scope of a contest can be narrow or broad. In narrow contests, the problem and requirements for the solution are clearly defined. Here, the sponsor lacks the resources or know-how to develop a good enough solution. Broad contests give fewer specifications and are more about searching for new markets or problems to solve (Adamczyk et al. 2012).

Contests benefit the sponsors and organisers by providing them intellectual property or public goods, access to lead users (Marchi et al. 2011) and opportunities to discover talented individuals to employ (Ågerfalk and Fitzgerald 2008). The design of the contests requires a suitable balance between incentives for rivalry and collaboration. Giving prizes only for the best solutions encourages rivalry, while using opportunities to learn as the primary motivator for participants can enhance collaboration. In their simplest form, contests consist of a problem definition, submissions by participants and an evaluation.

More complex contests include multiple evaluation phases and other activities during the process. Participants have an expectation of fair governance in contests, which requires stable and unambiguous rules. The evaluation of the contributions has a central role, and can be taken care of as an expert evaluation, peer review or popularity vote, or a combination of all three (Lampel et al. 2012).

Lampel et al. (2012) conclude that contests are an established innovation strategy, but it is still too early to determine their exact potential for innovation. Key challenges at the moment are the alignment of incentives with motives to foster collaboration. The high quantity and low quality of contributions is a common issue (Alexy et al. 2011).

3.2.8 How crowdsourcing innovations work

As discussed before, innovation is largely about identifying a problem, finding a solution and assessing the fit between them. Exactly when and why crowdsourcing is a suitable method for problem solving is still a matter of debate. However, its usefulness depends on the type of problem, the characteristics of the solution and the crowd (Afuah and Tucci 2012). Dynamics between users, process iterations, different selection mechanisms (Vuculescu and Bergenholz 2014) and award structures (Terwiesch and Xu 2008) can further impact the ability of the crowdsourcing system to find high-quality solutions. Where an organisation already understands the problem at hand and knows what solutions are likely to work, it is perhaps best to conduct the development project internally. When there is uncertainty about the problem (what the market needs) or solutions, then opening up the innovation process through crowdsourcing can be beneficial (Terwiesch and Xu 2008). The common approach is to use crowdsourcing to gather solution suggestions to more or less well-defined problems, but focusing on problems might be useful too. As research and development organisations are usually good at creating solutions, crowdsourced problems could be used as a starting point for a project pipeline, without challenging the professional status of employees (Cummings et al. 2012). Crowdsourcing can also give organisations access to lead users, that is, users who experience needs ahead of the main market (Marchi et al. 2011). It has been suggested that deploying crowdsourcing in business-to-business environments is more difficult than doing so in consumer contexts as there often is no direct contact with customers or end users (Simula and Vuori 2012). Finally, the nature of the problem and the means to solving it are interrelated. If the required solution is unlikely to come from a single actor, or if it requires building extensive new knowledge, it may be more useful to rely on collaborative communities or consortiums rather than crowdsourcing contests (Boudreau and Lakhani 2009).

Crowdsourcing innovations in contest format are a form of distributed or parallel idea generation. Here, idea refers to solution suggestion. Idea generation has been widely studied: a quick search on the term ‘idea generation’ brings up 1,367 search results on Scopus and 1,063 on the Web of Science Core Collection (as of 10 February 2015). Brainstorming approaches are a common way to generate ideas, and it is no surprise that the keyword is one of the most popular within the idea generation literature, measured in terms of the number of mentions and citations (Figure 3.3). Brainstorming has many potential benefits in the organisational context including supporting organisational memory of design solutions, impressing clients and creating income (Sutton and Hargadon 1996). However, standard group brainstorming may not be the best way to go if the quality of ideas is the main concern (Schirr 2012; Mullen et al. 1991). Instead, it has been suggested that a hybrid format, where people start by working individually and then as a group, produces the best results in terms of the number of different ideas generated and the quality of the best ideas generated (Girotra et al. 2010). In theory, crowdsourcing should therefore be a useful way to generate ideas, as it typically combines initial individual work with varying amounts of collaboration at the later stages. Indeed, it has been found, when compared to ideas generated by an internal development team that a crowdsourcing process can generate ideas that score higher in terms of novelty and customer benefit, although not in terms of feasibility (Poetz and Schreier 2012). Finding ways to improve the quality of the ideas generated in the crowdsourcing context is currently being investigated. For example, a crowd could be used to navigate large sets of analogies in search of new solutions (Yu et al. 2014).

Increased diversity among the problem solvers compared to internal sourcing could be one reason for the success of crowdsourcing (Terwiesch and Xu 2008). A large population of problem solvers includes people in the technical and social margins with different perspectives and heuristics. Such people have been shown to play an important role in successful problem solving (Jeppesen and Lakhani 2010).

With regard to the evaluation of solutions, crowdsourcing could offer a fast, cheap and effective way to generate relevance assessments (Alonso and Miszarro 2012). Crowdsourcing has been used for idea screening in stage-gate innovation processes (Onarheim and Christensen 2012) with promising results. Algorithmic approaches to improve the quality of the judgments have also been developed (Vuurens and de Vries 2012). Despite these developments, a final verdict on the usefulness of crowdsourced evaluations has not yet been made. Although there is a significant correlation between crowd evaluations and executive choices, evaluation bias has also been identified (Onarheim and Christensen 2012). The accuracy of evaluations could perhaps be improved by using multi-criteria scales instead of simple voting schemes or prediction markets, as the scales using multiple attributes have been found to significantly outperform them (Riedl et al. 2010; Blohm et al. 2011).

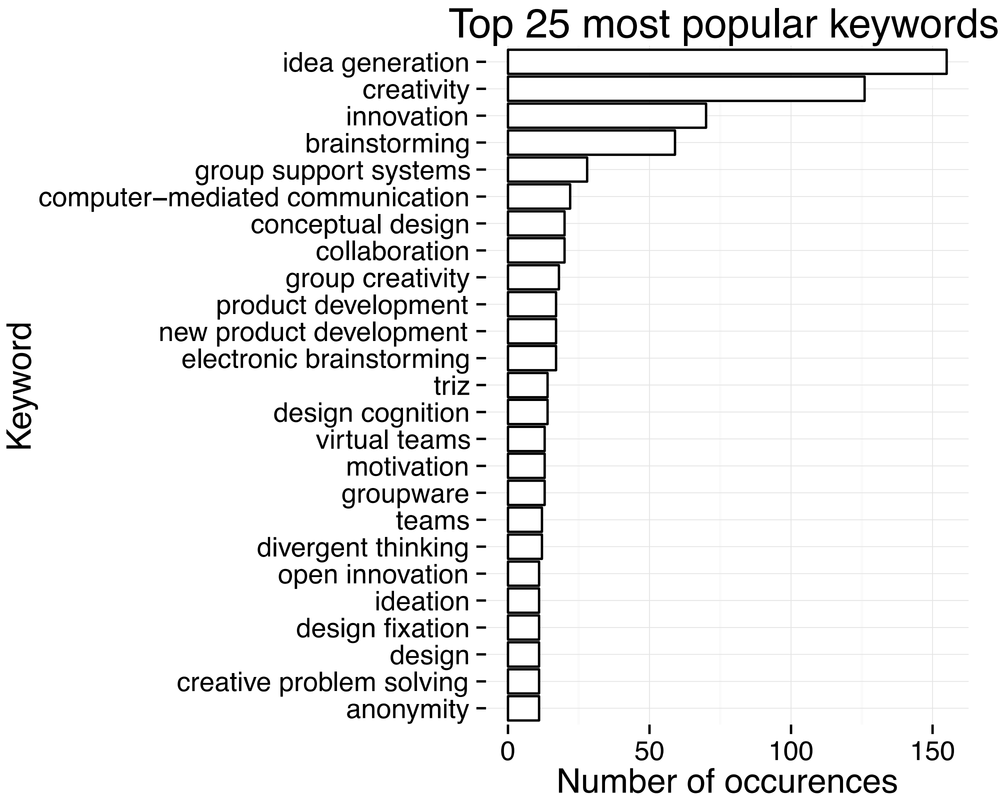


Figure 3.3: The most popular keywords in idea generation literature (Web of Science 2014).

Previous research demonstrates that crowdsourcing can be used to identify problems, possibly by tapping into lead users, to generate solution suggestions, and to evaluate these ideas. Using crowdsourcing explicitly to identify problems is not currently a common approach. Instead, the problem is usually defined by the crowdsourcing organisation. To be suitable for crowdsourcing, the problem should be relatively uncertain, that is, it should not be known in advance what kind of solutions will be the best. In this way, it is possible to take advantage of the diversity of solutions that crowdsourcing often provides. Limited scope is also important, so that individual actors can provide solutions. The creation of solutions through crowdsourcing is a form of distributed or parallel idea generation. In the light of existing research, crowdsourcing should be well suited for idea generation, as the applications often resemble the so called hybrid format of brainstorming, where participants first work alone, and then as a group. Idea generation in hybrid format has been found to provide good results both in terms of the quality and quantity of ideas. The use of crowdsourcing for evaluating ideas appears to be a promising approach, although the question is still open with regard to optimal evaluation mechanisms and their accuracy in practical contexts. In a less investigated research avenue, it has been suggested that collective intelligence might contribute to the effectiveness of crowdsourcing (Bonabeau 2009; Malone et al. 2010). Connecting large numbers of people around a shared problem-solving task might evoke phenomena that could be considered collective intelligence, a system that has qualitatively different properties than the individuals forming it (Heylighen 2013). Next we will look in more detail at what is already known about collective intelligence.

3.3 Collective intelligence

The study of collective intelligence in humans is a relatively new field, about which there are huge expectations including, for example, speculation on the emergence of the global brain (Heylighen 2011). New forms of collaboration made possible by the Internet, web 2.0 and social media add to the hype. It is, therefore, no wonder that interest in the field is rising. According to an often-cited definition[2], collective intelligence is a form of universal, distributed intelligence, which arises from the collaboration and competition of many individuals (Levy 1997). It is the general ability of a group to perform a wide variety of tasks (Woolley et al. 2010). The phenomenon is closely related to swarm intelligence, which means collective, largely self-organised behaviour emerging from swarms of social insects (Bonabeau and Meyer 2001). These terms have been used somewhat interchangeably; for example, Krause et al. (2009) define swarm intelligence as ‘when two or more individuals independently, or at least partially independently, acquire information and these different packages of information are combined and processed through social interaction, which provides a solution to a cognitive problem in a way that cannot be implemented by isolated individuals’. This is adopted here as a working definition of collective intelligence. I use swarm intelligence to refer to the emergent, collective behaviour of groups of cognitively simple agents[3] such as insects, robots and simulation algorithms. The term collective intelligence is reserved for phenomena involving agents with high cognitive capabilities, namely humans. This distinction is in line with the use of terminology in the literature. About 25% of the papers found on the Web of Knowledge using the keywords ‘collective intelligence’ discuss humans, whereas only 2% of the papers found with keywords ‘swarm intelligence’ do so.

Social psychology knows many examples of situations where group’s behaviour is not intelligent. For instance, groups can intensify opinions of their members (Myers 2012). This effect is known as group polarization: discussing a topic in a group can make the group members’ opinions more extreme than before discussion. For instance, if group members oppose war, they will oppose it even more after discussing with other like-minded people. The rise of the internet may have strengthened the effect. It is now easy for almost anyone to find support for their opinions in the chatrooms and forums on the internet (Gerstenfeld et al. 2003). Close-knit and cohesive groups with a strong leader and relative isolation from dissenting viewpoints are susceptible to groupthink. In high-stress

[2] Definition is cited 12 times in the analysed literature.
[3] Agent: one that acts or exerts power (Merriam-Webster 2015).

decision-making situations they may end up overestimating the group's abilities and moral justification, become closed-minded to new, contradictory information, and, due to conformity pressures and self-censorship, develop an illusion of unanimity. Groupthink is similar to confirmation bias affecting individuals, which refers to seeking and interpreting information in ways that are partial to existing beliefs and expectations (Nickerson 1998). Vast empirical evidence suggests that the confirmation bias is common, strong and appears in many guises. For example, confirmation bias leads individuals to ignore information that goes against their original hypothesis, look primarily for positive cases, and ignore counterexamples and alternative hypotheses. As a result opinions once formed tend to last, and attention sifts in defending them. Confirmation bias has been found to increase polarisation of groups: both parties interpret a new piece of information as supporting their side (Nickerson 1998). In addition to getting locked in a viewpoint, groups can also be led astray by social influence bias. Large-scale web-based experiments have shown that aggregated opinions of others can influence decision making of individuals. Prior ratings create significant bias in individual rating behaviour (Muchnik et al. 2013; Salganik and Watts 2008; 2009). This bias can influences rating dynamics in systems designed to harness collective intelligence (Muchnik et al. 2013). These and other well-known examples of "madness of crowds" are not in the focus of this research. Instead, the interest is in the situations where interactions increase the cognitive skills and decision making accuracy of a group or crowd.

Approaches to studying collective intelligence have been diverse, from the purely theoretical (Szuba 2001; 2002) and conceptual (Luo et al. 2009) to simulations (Bosse et al. 2006), case studies (Gruber 2007), experiments (Woolley et al. 2010) and systems design (Vanderhaeghen and Fettke 2010). The field is also multidisciplinary as it is related to psychology (Woodley and Bell 2011), complexity sciences (Schut 2010), cognitive studies (Trianni et al. 2011), biology (Bonabeau and Meyer 2001), computer sciences, semantics (Levy 2010) and social media (Shimazu and Koike 2007). At the moment, there is no theory capable of explaining how collective intelligence actually works (Schut 2010). Despite some efforts (e.g., Luo et al. 2009; Gan et al. 2007; Malone et al. 2010), generally accepted frameworks for studying collective intelligence in humans do not exist. As a result, the field is at risk of fragmentation. Although a certain amount of diversity is probably good for the advancement of a scientific field (Woolley and Fuchs 2011), the lack of an overarching structure can make the field appear confusing. In addition, it is challenging to tie the efforts of different disciplines together in a coherent way. Due to the lack of a common framework, it is not possible to assess what is already known. It is challenging for researchers from different disciplines to be aware of advancements in other fields, where the same concepts may be named differently.

Here, the focus is on the question of what the scientific community means by the notion of collective intelligence in the human context. The objective is to review the current literature, identify relevant themes and form a conceptual framework for studying the phenomenon. The scope of the review is limited to literature discussing collective intelligence in humans. The limitation is based on the assumption that rich and complex cognitive and psychological behaviour sets humans apart from insects, algorithms and robots. Furthermore, the development of human intelligence resulted in part from evolutionary pressures to navigate in social situations to one's own benefit (Geary 2005). Such behaviour could plausibly undermine the collective performance of groups, or at least make it significantly different from situations where individual motivations are mostly aligned. At this point it would be premature to try to combine phenomena from different contexts under one framework without first understanding each context separately.

The selection of literature for this review follows the approach of Zott et al. (2011). A keyword search was conducted on the Web of Knowledge on 7 July 2011 using the keywords 'collective intelligence' and 'swarm intelligence'. The searches produced 405 and 646 results, respectively. In addition, all issues of the journals *Swarm Intelligence* and the *International Journal of Swarm Intelligence* were reviewed for suitable articles. A cursory analysis was performed by reading through the titles and abstracts. The following criteria were used to select the papers for review: 1) the paper discusses collective intelligence in the human context; 2) the publication in which the paper is published is listed on the Web of Knowledge; and 3) the paper makes a non-trivial contribution to the discussion about collective intelligence (i.e., it involves more than a couple of mentions of the term). Using these criteria, 41 papers were selected. The papers are marked with an asterisk (*) in the References section. The purpose was not to cover everything that has been written about the topic, but to review a representative sample of papers to gain an adequate understanding of the relevant themes of collective intelligence on humans. The papers were read thoroughly and definitions of collective intelligence and related themes were collected. The main contributions to collective intelligence research were identified. Similar definitions and themes were grouped together and the resulting categories were named as seemed appropriate. Sticky notes were used to make the process visual and thus help the recognition of interesting patterns in the data. Additional references were gathered and further limited literature searches were performed to fill in the gaps (e.g., definitions of self-organisation, trust and emotional intelligence) and thus to provide a more complete view on what is already known about collective intelligence in humans.

The grouping of themes and definitions revealed a pattern in the literature. The discussion of collective intelligence in humans appears to be divided into three levels of abstraction: the micro level, the macro level and the level of emergence.

Table 3.3 identifies the themes in the literature and groups them under the three levels of abstraction. Examples of papers discussing these themes are also given.

3.3.1 The micro-level: Enabling factors of collective intelligence

At the micro-level, collective intelligence is a combination of psychological, cognitive and behavioral elements. Pentland (2007) argues that humans should firstly be viewed as social animals and only secondarily as individuals. According to his research with the so-called Socioscope, human behavior is largely predictable, non-linguistic signal-response behavior. The immersion of self in a social network is a typical human condition and our unconscious ability to read and display social signals allows smooth coordination within the network. Pentland suggests that important parts of human intelligence could thus reside in network properties. This might just be the case, as Woolley et al. (2010) found evidence of the existence of a single dominant collective intelligence factor, c, underlying group performance. In their experiments, c explained 30-40 % of group performance and was found to depend on the composition of the group (e.g. average intelligence) and emergent factors resulting from interaction of group members, such as conversational turn-taking. Furthermore, c is positively correlated with social sensitivity and the proportion of females in the group, but the influence of females is probably mediated by their better average social sensitivity (Woolley et al. 2010). Many open questions remain regarding the nature of c. Woodley and Bell (2011) suggest that c could actually be largely a manifestation of the General Factor of Personality (Just 2011) at a group level.

Other relevant themes are trust (Scarlat and Maries 2009, Bosse et al. 2006) and attention (Zembylas and Vrasidas 2005, Gruber 2007, Trianni et al. 2011). A certain level of trust is a precondition for cooperation. Attention is used as an implicit measurement of value in many contemporary web applications, such as YouTube (view count) and Twitter (re-tweets).

Table 3.3: A list of themes related to collective intelligence in humans categorised under three levels of abstraction

Level	Theme	Definition	Examples of papers from the sample
Micro	Humans as social animals	Viewing humans as social animals: immersion of self in a social network, a typical human condition	Pentland 2006; 2007
	Intelligence	The intelligence of individual human beings, often measured with the g-factor	Woolley et al. 2010
	Personal interaction capabilities	The factors affecting a person's ability to interact with other human beings, such as emotional intelligence (Cherniss 2010), social sensitivity (Woolley et al. 2010) and the general factor of personality (Just 2011)	Woolley et al. 2010; Woodley and Bell 2011
	Trust	An actor's expectation of the other party's competence and goodwill (Blomqvist 1997)	Bosse et al. 2006; Scarlat and Maries 2009
	Motivation	The factors influencing the interest to participate in communities or to contribute to collective effort	Franck 2002; Rasmussen et al. 2003; Bonabeau 2009; Lykourentzou et al. 2010; Brabham 2010; Malone et al. 2010
	Attention	The commitment of cognitive resources	Zembylas and Vrasidas 2005; Zettsu and Kiyoki 2006; Gruber 2008; Trianni et al. 2011
	Communities	Real and virtual communities, such as communities of practice and online social networks (Cachia et al. 2007) and brand communities (Brabham 2010)	Coe et al. 2001; Cachia et al. 2007; Chen 2007; Lykourentzou et al. 2010; Brabham 2010
Emergence	Complex adaptive systems	Systems that show adaptivity, self-organisation and emergence (Ottino 2004)	Komninos 2004; Chen 2007; Luo et al. 2009; Schut 2010; Trianni et al. 2011
	Self-organisation	The emergence of order at the system level without central control, solely due to local interactions of the system's components (Kauffman 1993)	Bonabeau and Meyer 2001; Franck 2002; Rasmussen et al. 2003; Wu and Aberer 2003; Luo et al. 2009; Krause et al. 2009; Schut 2010; Trianni et al. 2011
	Emergence	A rise of system-level properties that are not present in its components; 'the whole is more than the sum of its parts' (Damper 2000)	Rasmussen et al. 2003; Chen 2007; Cachia et al. 2007; Luo et al. 2009; Schut 2010; Lee and Chang 2010; Woolley et al. 2010; Trianni et al. 2011
	Swarm intelligence	The study of cognitively (relatively) simple entities, whose collective behaviour is intelligent	Bonabeau and Meyer 2001; Wu and Aberer 2003; Krause et al. 2009; Luo et al. 2009; Trianni et al. 2011
	Stigmergy	A mechanism of indirect coordination, originally describing the nest-building behaviour of termites (Theraulaz and Bonabeau 1999)	Bosse et al. 2006
	Distributed memory	The shared, often external, dynamic memory system that performs parts of agents' cognitive processes (Bosse et al. 2006)	Bosse et al. 2006; Scarlat and Maries 2009; Gregg 2009; Luo et al. 2009; Levy 2010; Trianni et al. 2011
Macro	Decision making	The process of making decisions, both individually and in groups	Pentland 2006; Bonabeau 2009; Malone et al. 2010; Gregg 2010; Krause et al. 2011
	Wisdom of crowds	Under certain conditions, groups can be more intelligent than the smartest individuals in them; a collective estimate can be accurate, even if individual estimations are not (Surowiecki 2005)	Chen 2007; Pentland 2007; Nguyen 2008; Krause et al. 2009; Brabham 2009; Lykourentzou et al. 2010; Leimeister 2010; Lee and Chang 2010; Brabham 2010; Lorenz et al. 2011
	Aggregation	The combination of individual pieces of information to form a synthesis or collective estimation	Pentland 2007; Bothos et al. 2010; Krause et al. 2011
	Bias	The tendency of individuals and groups to make systematic errors in decision-making situations	Cachia et al. 2007; Gregg 2009; Lee and Chang 2010; Krause et al. 2011
	Diversity	The differences in demographic, educational and cultural backgrounds and the ways that people represent and solve problems (Hong and Page 2004)	Bonabeau and Meyer 2001; Bonabeau 2009; Brabham 2010; Krause et al. 2011
	Independence	The decision of an individual is not influenced by the decisions of other individuals	Lorenz et al. 2011

3.3.2 The macro-level: Output of the system and wisdom of crowds

At the macro level, collective intelligence becomes a statistical phenomenon, at least in the case of the wisdom of crowds effect (Lorenz et al. 2011). The term wisdom of crowds, coined by Surowiecki (2005), describes a phenomenon where, under certain conditions, large groups can achieve better results than any single individual in the group. For example, the average of several individuals' estimates can be accurate even if individual estimations are not. The wisdom of crowds effect is claimed to be based on diversity, independence and aggregation (Surowiecki 2005).

Diversity in groups of people usually refers to differences in demographic, educational and cultural backgrounds and differences in the ways that people represent and solve problems (Hong and Page 2004). Both a simulation model (Hong and Page 2004) and an experiment with humans (Krause et al. 2011) have shown that under certain conditions groups of diverse problem solvers can outperform groups of high-ability problem solvers. In addition, the best problem solvers were biased in their estimations, while the group, as a whole, was accurate (Krause et al. 2011). A question remains as to whether this finding was unusual or to be expected.

Independence means that the estimations of one individual are not influenced by the estimations of other individuals. Lorenz et al. (2011) have shown that even minor social interaction can undermine the wisdom of crowds, which happens through three effects. The social influence effect reduces the diversity of a group without increasing its accuracy. The range reduction effect causes the correct value to become less central in the distribution of evaluations, thus delivering a false hint regarding the location of the truth. The confidence effect is a psychological result of the two statistical effects described above. It increases individuals' confidence in their estimations even though collective accuracy has not improved. Lorenz et al. (2011) propose that these effects occur at a certain range of difficulty for decision making and confidence in decision makers.

Aggregation refers to mechanisms for pooling and processing individual estimations to a collective estimation. While simple averaging might be the most common method of aggregation, it is not always the most suitable. In many cases, other statistical aggregate measures should be considered (Lorenz et al. 2011). The rise of the Internet has also made it possible to develop new aggregation methods, such as information aggregation or prediction markets (Bothos et al. 2009), social tagging or folksonomies (Gruber 2007; Zettsu and Kiyoki 2006) and data visualisation (Chen 2007).

3.3.3 The level of emergence: From local to global

The level of emergence resides between the micro level and the macro level and deals with the question of how system behaviour on the macro level emerges from interactions of individuals at the micro level. A common approach to explaining how collective intelligence emerges from individual interactions as a statistical or probabilistic phenomenon is to use the theories of complex adaptive systems. Complex adaptive systems are characterised by adaptivity, self-organisation and emergence (Ottino 2004). Adaptivity means the ability of a system, or its components, to change itself according to changes in the environment (Schut 2010). Self-organisation means the emergence of order at the system level without central control, due solely to local interactions of the system's components. The basic ingredients of self-organisation include positive and negative feedback loops, randomness and multiple interactions (Bonabeau 1999). A simple definition for emergence is that 'the whole is more than the sum of its parts' (Damper 2000). Extending from these premises, Schut (2010) proposes three enabling properties and five defining properties for collective intelligence systems. The existence of adaptivity, interaction and rules executed at a local level make it possible for collective intelligence to emerge from a system. If the system shows a distinction between global and local, randomness, emergence, redundancy and robustness, it is a collective intelligence system. Group memory (Trianni et al. 2011), a shared extended mind (Bosse et al. 2006) and other similar concepts are also relevant to the emergence of collective intelligence. Bosse et al. (2006) give the following criteria for a shared extended mind:

1. The environment participates in the agents' mental processes.
2. The agents' internal mental processes are simplified.
3. The agents have a more intensive interaction with the world.
4. The agents depend on the external world in the sense that they delegate some of their mental representations and capabilities to it.

A shared extended mind thus works as a dynamic short-term memory that allows the coordination and collaboration of individual components of the complex adaptive system. Notably, the components creating the shared extended mind need not be aware of it, nor benefit from its creation (Bosse et al. 2006).

The literature provides many examples of swarm intelligence systems that display the characteristics of complex adaptive systems and a shared extended mind. The behaviour of social insects is maybe the most classical example. The foraging of ants (Camazine et al. 2001), the nest-site selection of honeybees (Seeley and Buhrman 1999) and the nest building of termites (Turner 2011) all use some form of distributed memory and show emergent, adaptive behaviour as a result of self-organisation. The features of complex adaptive systems have also been considered relevant in the context of human collective intelligence (Komninos 2004; Luo et al; 2009, Chen 2007). A number of these features have been demonstrated in case studies (Wu and Aberer; 2003, Bonabeau 2009; Lykourentzou et al. 2010). The Internet as a shared memory of humankind has been mentioned repeatedly (e.g., Levy 2010; Luo et al. 2009; Heylighen 1999).

3.3.4 The theoretical framework of collective intelligence

Using insights from the current literature, it is now possible to construct a theoretical framework to guide the research. Theoretical or conceptual frameworks explain what is to be studied: the key variables and the presumed relationships among them (Miles and

Huberman 1994). They can be simple or complicated and their development can rely on existing theories or common sense. Graphical format for conceptual framework is preferable, but narrative form is also possible. The main benefit of building frameworks is that they force the researcher to clarify and make explicit the assumptions about the studied phenomena and be selective about the most important variables and relationships (Miles and Huberman 1994). Table 3.4 shows the relationships between themes from literature and elements of theoretical framework of collective intelligence.

Table 3.4: Relationships between themes from literature and theoretical framework of collective intelligence.

Level	**Theme**	**Element of theoretical framework**	**Notes**
Micro	All	Human capabilities for interaction	
Emergence	Complex adaptive systems	Agents, activities, feedback, emergence	
	Self-organisation	Agents, activities, feedback	
	Emergence	Emergence	
	Swarm intelligence	-	The study of cognitively (relatively) simple entities, whose collective behaviour is intelligent.
	Stigmergy	Agents, activities, feedback, distributed memory	Facilitates swarm intelligence.
	Distributed memory	Distributed memory	
Macro	Decision making	Output	
	Wisdom of crowds	Output	
	Aggregation	-	A requirement for wisdom of crowds.
	Bias	-	A characteristic of decision making.
	Diversity	-	A requirement for wisdom of crowds.
	Independence	-	A requirement for wisdom of crowds.

As discussed above, collective intelligence refers to phenomena where the intelligence of a group can be considered to be at least partially independent of and usually greater than the intelligence of individuals forming the group. This study uses the framework from the literature in a simplified form as shown in Figure 3.4. The idea is that by collecting data about the elements presented in the figure, the study will cover all relevant aspects in order to develop an understanding of collective intelligence.

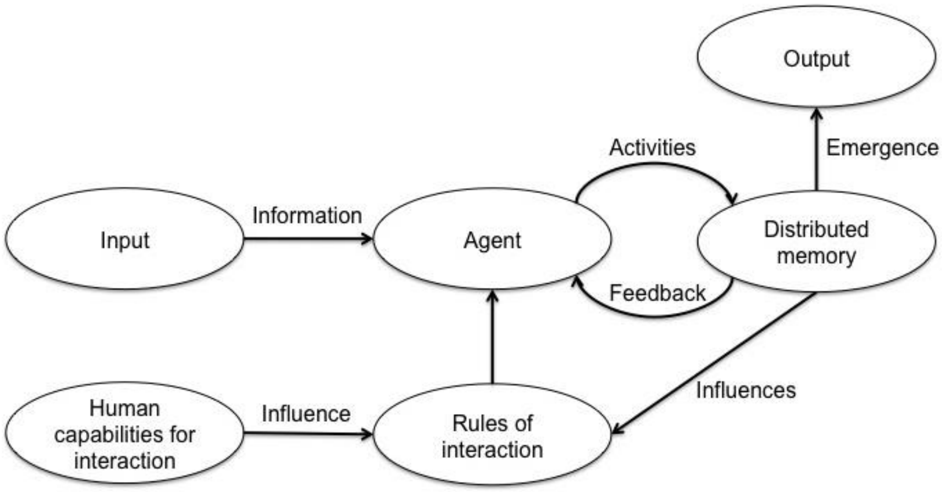


Figure 3.4: The theoretical framework of collective intelligence used to guide data collection

According to the framework, human capabilities for interaction, such as intelligence, trust, motivation, and other psychological and cultural factors, together with environmental constraints, create the rules of interaction. Inputs to the system arrive through cognitive agents. An agent processes and integrates information from the outside and feedback from the distributed memory. Actions are performed according to more or less strict rules. The distributed memory is the shared environment of the agents, which stores the information they create. Actions can also change the state of distributed memory. Changes to memory are fed back to the agent, and may also change the environmental constraints. Out of the multiple interactions between agents and distributed memory emerges the output of the system. Agents, their rules of interaction, distributed memory and environmental constraints form a complex adaptive system, which reacts to information from outside. The output is the emergent property of the system and may demonstrate the wisdom of crowds: the decisions made by the system as a whole may be of better quality than those that individuals can produce on their own. These high quality decisions result from diversity, independence and information aggregation.

Based on the framework, it is proposed, first, that the micro-level features of human beings, such as intelligence, trust and motivation, are the enabling factors of collective intelligence. They provide the rules according to which individuals act. Micro-level features set humans apart from other collective intelligence systems; for example, motivation does not have to be taken into account when designing robots or algorithms. Secondly, individuals interacting with each other form a complex adaptive system, which shows self-organisation and emergence. Distributed memory facilitates communication and coordination between individuals. A comparison of collective intelligence in humans with examples of swarm intelligence in other contexts might be most fruitful at this level of abstraction. Thirdly, the global behaviour of the complex adaptive system is probabilistic by nature. At this level, diversity, independence and mechanisms of information aggregation are important features of the system. Measuring these features should help predict the global performance of the system as a whole. Combining different approaches to the study of collective intelligence in humans seems possible despite the multidisciplinary nature of the phenomenon. The three levels of abstraction offer different lenses through which collective intelligence can be viewed. The viewpoints complement each other to provide a fuller picture of the phenomenon.

3.3.5 Collective intelligence genome

Next other relevant frameworks for the study of collective intelligence are reviewed. The frameworks are called the genome of collective intelligence, collective intelligence systems, and distributed cognition. Later in the study they are used to provide different perspectives or "theoretical lenses" for analysing the results.

Genome of collective intelligence is a framework for classifying the building blocks of collective intelligence systems developed by Malone et al. (2010). Taking an analogy from biology, it strives to describe various collective intelligence systems as combinations of a relatively small set of building blocks. These building blocks are classified using four questions: what is being done (goal), who is doing it (staffing), why are they doing it (incentives), and how is it being done (structure/process). A basic organisational building block or gene is defined as a particular answer to one of the four key questions; the combination of genes forming the description of a system is called the collective intelligence genome of the system. Classification of genes identified so far is presented in Table 3.5.

According to Malone et al. (2010), many of the organisational tasks encountered in collective intelligence systems are either about creating something new, such as a piece of software or a t-shirt design, or making a decision, evaluation or selection, for instance, on which t-shirt design to print or what code to include in the software project. As an answer to the question who undertakes the activity, there are also two basic alternatives: hierarchy or crowd. Hierarchy refers to hierarchical organisations, where someone in a position of authority assigns a task to a particular actor to complete. Crowd on the other hand refers to situations where the activities can be undertaken by anyone in a relatively large and undefined group of people. This reliance on crowd at least in some part of the activities is a central feature of web-based collective intelligence systems.

The question why people take part in an activity deals with the motivation. The genome of collective intelligence framework offers only a very simplified overview of the possibilities with three basic alternatives: money, love and glory. These deal with financial gain, intrinsic enjoyment and desire for recognition, respectively. This classification is somewhat vague and is not directly in line with theories on motivation. Therefore, this part of the framework is replaced with an even simpler classification of intrinsic and extrinsic motivational factors (Deci et al. 1999). Assessments of motivation should nevertheless be taken only as a rough approximation as to which of the two possibilities is more important.

Table 3.5 Building blocks or genes of collective intelligence systems (adapted from Malone et al. 2010)

Question	**Gene**	**Useful when (or definition)**
Who	Crowd	Undefined group of people
	Hierarchy	Hierarchical organisation
Why	Extrinsic motivation	An activity is performed to attain a certain goal
	Intrinsic motivation	An activity is performed for the sake of interest or enjoyment of the task itself
How - create	Collection	Crowd members create individual items
	Contest	Collection, where some of the items are selected as winners
	Collaboration	Crowd members work on a common project with important interdependencies between the contributions
How - decide	Group decision	Decisions of crowd members are aggregated to form a single decision that applies to all crowd members
	Voting	Group decision where individual opinions are aggregated by a vote count
	Averaging	Group decision where individual opinions are aggregated by averaging scores
	Consensus	Group decision that is reached through discussion
	Prediction market	Speculative artificial market, where market prices are interpreted as predictions of probabilities of events
	Individual decisions	Decisions of crowd members are not combined
	Market	Some kind of formal exchange is involved in the decision
	Social network	Crowd members form a network of relationships, allowing development of trust and information transfer between the crowd members, for example

The final category of building blocks deals with the question of how task is carried out. In hierarchies, the answer is typically a description of organisational processes. When dealing with crowds, the key determinant is whether the participants make the contributions independently of each other. The building blocks in this category interact with those describing what is done. There are four alternatives in total: creation can take place as a collection or in collaboration; and decisions can be individual or group decisions. A collection is simply a collection of items created by individual members of the crowd. Contests are a typical example of systems that produce collections. In contests, one or more items the crowd has provided are selected as winners and the corresponding crowd members receive prizes. In collaboration, members of the crowd work together on a common project and the contributions of individual members have important dependencies related to the contributions of others. A group decision takes place when the members of a crowd make a decision that applies to the crowd as a whole, for instance, by voting or aggregating the opinions in some other way or through consensus. In individual decisions, the decisions of the crowd members are not aggregated together.

Markets and social networks are given as examples of places where individual decisions take place.

3.3.6 Collective intelligence systems

As discussed in the Introduction, social insects show interesting group behaviour where relatively simple agent interactions result in the emergence of much more complex problem-solving capabilities, such as regulating foraging, selecting nest sites, and building nests (Gordon et al. 2008; Visscher 2007; Turner 2011). This collective, largely self-organised behaviour is called swarm intelligence (Bonabeau and Meyer 2001). Schut (2010) gives a framework for evaluating these kinds of collective intelligence systems. There are three enabling and five defining properties of these systems. If enabling properties are observed, the system can be a collective intelligence system. If defining properties are also in place, the system can be called a collective intelligence system. Table 3.6 summarises these properties. The collective intelligence systems framework is in line with the theoretical framework of collective intelligence derived from the literature: the emphasis is on the emergence of macro-level (global) behaviour from micro-level (local) interactions.

Table 3.6: Enabling and defining properties of collective intelligence systems (Schut 2010)

Property	Definition
Enabling properties	
Adaptivity	Changing one's structure to fit the environment: individuals, rules or the system.
Interaction	Individual behaviours and interaction between individuals
Rules	Implications between inputs and outputs
Defining properties	
Global-local	Individuals in the system vs. the system as a whole
Randomness	Elements of randomness in the system
Emergence	'The whole is more than the sum of the parts'
Redundancy	The same information represented in many places
Robustness	Even if some parts fail, the system stays functional

The enabling properties of collective intelligence systems are mainly concerned with the local level or the agents of the system. These are adaptivity, interaction and rules. Adaptivity means changing one's structure to fit the demands of the environment. These changes can take place at the individual or system level. Changes to individual behaviour imply changes o at the system level but not necessarily the other way around. For instance, ant colonies can adapt to forage efficiently from differently distributed food sources even when the behaviour of individual ants stays the same (Camazine et al. 2001). The different behaviour at the system level results from the interactions between large numbers of participating individuals. When dealing with complex systems, it is not possible to analyse the behaviour of the system by considering a part of it in isolation. Both individual behaviours and the interactions have to be taken into account. The interactions are related to the concept of information flows, where information from one part of the system can be propagated to other parts of the system and cause major changes in system behaviour. Finally, the rules are a way of describing the behaviour. There are implications between inputs and outputs, or observations and actions, either at the individual or the system level. Rules can be presented in multiple forms, for instance, as simple if/then rules. With regard to human beings, the rules are closely related to behaviourist views on human psychology.

The defining properties of collective intelligence systems deal with the global level of the system. The first property is the distinction between local and global aggregation levels. The local level is concerned with the individuals in the system, whereas the global level is about the system as a whole. This distinction is necessary, for example, for positioning the concepts of adaptivity and emergence. As mentioned, adaptivity can take place either at the local or the global level. Emergence on the other hand is concerned with moving from the local to the global level in a complex system: at some point along the way, qualitatively different behaviour emerges and the system becomes 'more than the sum of its parts'. Emergence is the most challenging property of collective intelligence systems. Wolf and Holvoet offer a working definition for the property: 'a system exhibits emergence when there are coherent emergents at the macro-level that dynamically arise from the interactions between the parts at the micro-level. Such emergents are novel with respect to the individual parts of the system' (De Wolf and Holvoet 2005). Complex systems typically have some elements of randomness as a result of which the systems show critical behaviour: they exist at the edge of chaos. Most of the time, the systems are in balance or in a state where small disturbances have no effect and responses are proportional to the impact. Only dramatic disturbances can cause dramatic changes. However, certain external events can push the system out of balance and over the edge into chaos. In such cases, smaller impacts can trigger large changes at the system level. The last two properties, redundancy and robustness, are interrelated. Redundancy means that same information is presented in multiple locations in the system. In this way, the system can lose some of its parts without losing its functionality. The system is robust against damage.

3.3.7 **Distributed cognition**

The cognitive sciences are very suitable for studying thinking, and especially distributed cognition theory offers a useful framework for addressing the questions of this thesis. The assumption is that cognitive properties and processes are causes for observed behaviour. However, knowing the properties of individual agents is not always enough for understanding the performance of a system. Whereas the cognitive sciences usually take individuals as the unit of analysis, distributed cognition broadens the perspective by considering the information processing of groups of interacting individuals in distributed socio-technical systems. The cognitive properties of distributed systems can be radically different from the cognitive properties of the individuals involved in them. Furthermore, the application of the classical cognitive science viewpoint to larger, distributed cognitive systems with little modification is feasible, for instance, in the contexts of airline cockpits (Hutchins and Klausen 1995, Hutchins 1995) and traditional aircraft carrier navigation

(Hutchins 1996). In these systems, cognitive labour is distributed as a single individual cannot typically do the required jobs alone. In addition, individual agents form only a part of the system, and other parts of the systems, such as technical devices, can also do important cognitive work. For instance, a pen and paper can be used to store information. A map, compass and other navigation equipment essentially take care of complex calculations when defining a vessel's position on the ocean (Hutchins 1996). A fundamental question concerns how cognitive processes are implemented in a group of individuals (Hollan et al. 2000). Distributed cognitive systems process information by propagating a representational state across various representational media (Hutchins and Klausen 1995). The representational media can reside inside individuals, between them (as speech, for instance) or in physical structures and artefacts. Each medium has its own properties that constrain the types of cognitive processes that can be used with the medium. The cognitive properties of a system are partially determined by the cognitive properties of individuals, representational media and properties, the organisation of the media and access to relevant information across individuals. The system level properties emerge from the interactions of individuals and media within the system.

In contrast to traditional cognitive sciences, where most of the time what goes on in an individual's mind can only be inferred, in distributed cognitive systems parts of the information processing can be observed directly. The movement of information through a system can have important consequences, for instance, by organising and coordinating the behaviour of individuals. Although it is not possible to know with certainty where the information will go, tracking possible paths is often feasible, by finding out who could have had access to information and when. An additional challenge to this task is that not all pathways are designed. Some may be unintended. Often it is possible to determine from other evidence where the information actually went, once the possible pathways are known. Computation is performed by propagation and transformation of representational states throughout the system (Hutchins and Klausen 1995).

4 Case descriptions

In this chapter, the three cases are described in detail. Case descriptions are an important analytical tool in qualitative analysis (Miles and Huberman 1994). They condense, filter, and summarise the large amounts of disparate raw data into a coherent and more accessible format. The case descriptions form a basis for further qualitative analyses. The information gleaned from the descriptions is later reorganised using several theoretical lenses to drill down into details relevant to collective intelligence. As qualitative analysis is mostly carried out with words, special attention is paid to keeping the language of case descriptions as close to the original as possible. Readability is slightly compromised to maintain the chain of evidence intact. Descriptions of user experience are based on experiences of the researcher and descriptions of other users (for example forum messages) included in the collected data.

4.1 OpenIDEO

OpenIDEO defines itself as follows:

> “OpenIDEO is a place where people design better, together for social good. It's an online platform for creative thinkers: the veteran designer and the new guy who just signed on, the critic and the MBA, the active participant and the curious lurker. Together, this makes up the creative guts of OpenIDEO” (OpenIDEO 2012).

OpenIDEO is a website hosted by design and innovation firm IDEO, a renowned global design company with a human-centred approach to design. The website is dedicated to designing social innovations collaboratively and including a broader range of people in the design process. The activities on the site are organised around challenges. The challenges are difficult design tasks, which are usually related to some large and complex environmental or societal issue, such as food production, health care, or unemployment. Organisations and individuals can sponsor a challenge. Figure 4.1 depicts a screenshot from the OpenIDEO website.

On the OpenIDEO site, users can submit inspirations or concepts, update their own concepts, evaluate concepts, and comment and applaud blog posts, inspirations and concepts. Commenting and applauding are possible in every phase and even after the challenge has ended. Other activities are possible only for a limited time during the corresponding phase.

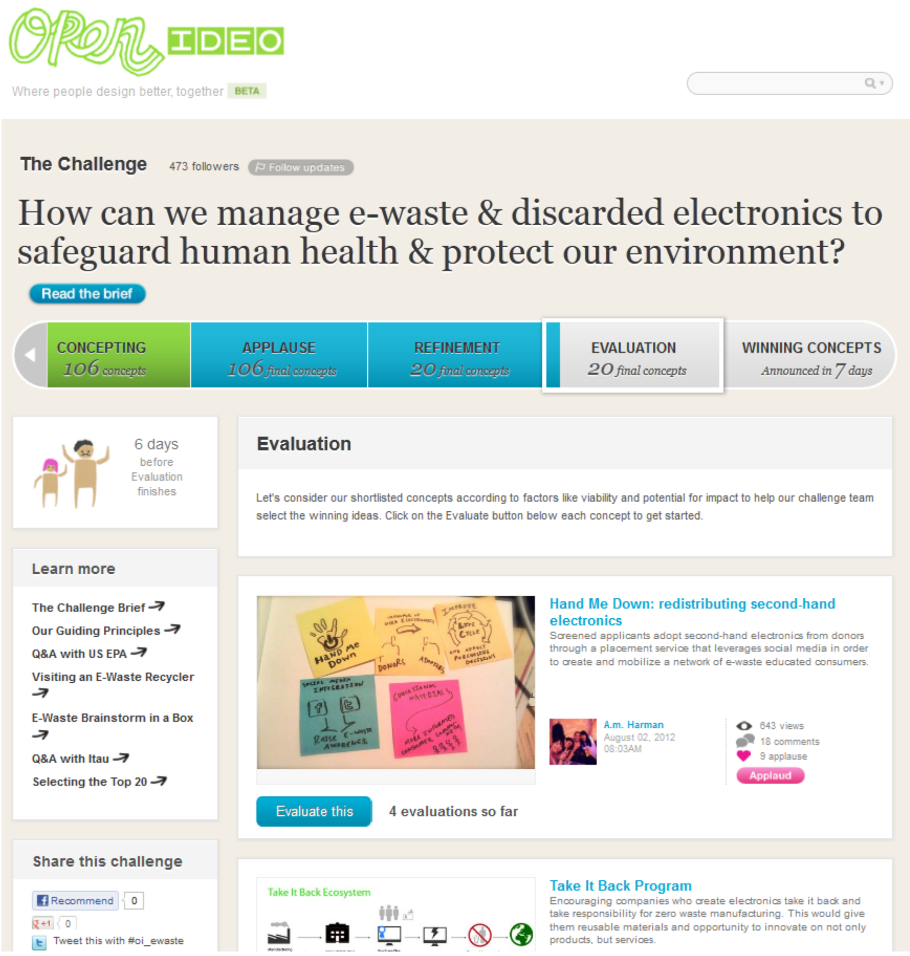


Figure 4.1: The OpenIDEO website

4.1.1 Rules at OpenIDEO

OpenIDEO is an online platform for creative thinkers who care about social good. It seeks to be inclusive, community-centred, collaborative, optimistic, and always in beta (that is, never cast in stone). Organisations and individuals can sponsor a challenge for social or environmental good. OpenIDEO's approach to managing the innovation process is soft and indirect, and seems to be based on creating a shared culture. Instead of direct tasks and explicit rules, the tasks are given indirectly and rules are enforced gently but firmly in the discussions that take place on the site. The approach used to create innovations has been termed collaborative competition: although there are winners, collaboration is encouraged at every turn. Apart from the community's appreciation, winners get no rewards. According to the principles of OpenIDEO, the site is an online platform for creative thinkers who care about social good. This is the place where translation of stellar skills into real world action is celebrated. Social impact is the big focus of the collaborative community at OpenIDEO. The organisation favours the transformation of ideas as a way to make an impact. The basic principles of the OpenIDEO community are shown in Figure 4.2.

Principles of the OpenIDEO Community

Principle #1: Inclusive
Recognize and enable all levels of participation from different disciplines. It's about allowing anyone to contribute to the creative process. Whether it's a great insight, a beautiful sketch, an encouraging build, or a few words of praise, the platform allows everyone to take part in and feel as if they are a part of the process.

Principle #2: Community-centered
Remember the core strengths of the community and play to them. OpenIDEO is meant to foster a strong, vibrant, lively community that thrives on inspiration — and that we all trust will make a difference. Focusing on this community and its activities is essential.

Principle #3: Collaborative
Promote teamwork among individuals and teams by recognizing the many roles that are crucial to each step of the design process. Always choose collaboration over competition, and create an atmosphere for building on one another's ideas.

Principle #4: Optimistic
Stay optimistic! You never know when a wild idea might enable others to get closer to a viable solution.

Principle #5: Always in Beta
Design for continuous improvement and iteration and scale deliberately. That goes for the community, the platform, and these principles. To this end, please email your suggestions for making OpenIDEO better.

Figure 4.2: Principles of the OpenIDEO Community (OpenIDEO 2012)

The legal terms and conditions seem to follow the standard approach for websites. Perhaps worth noting is the fact that there appear to be two sets of terms and conditions: the general terms and conditions of the site and challenge-specific terms and conditions, which can overrule the general terms. For example, the general terms and conditions state that the personal information of users will never be shared, unless stated otherwise in the challenge terms and conditions. The participants own the intellectual property they have created on the site, but license it to the challenge sponsor. Concepts can also be freely used, shared and remixed, similar to the terms of creative commons licenses. Users should acknowledge original sources by using quotation marks and giving links to sources. Inspirations should be linked to concepts by using the Build this feature provided on the site.

4.1.2 Feedback at OpenIDEO

Feedback is immensely important for the functioning of the OpenIDEO site. Although feedback practices are rarely mentioned explicitly, the amount of feedback on the site is significant. There are two general sources of feedback: the site itself and its official representatives, and the other users. Feedback is given through written comments, blog posts and by displaying the numbers of comments and counts of applause each contribution has accrued. Several flavours of feedback can be identified. A typical comment could look like the following generic example: ‘Great concept! I like how it combines the ideas suggested by Tom and Jerry. Have you thought about how this could be used if electricity is not available? Thanks for sharing your thoughts!’

Instant feedback from virtual collaborators and, as one participant on the site mentioned, knowing that someone is looking over your shoulder are important motivational factors. In my personal experience, just a few positive comments and tips made a big difference. The comments I received for the concept I submitted were thorough and the user had clearly put some thought and effort into them. In response to the comments, I ended up spending a couple of hours refining the concept. Without the feedback I definitely would not have worked on the concept on a Sunday. In contrast, my other concept did not generate similar feedback, and I never returned to work on it. I also developed the habit of checking the number of views and comments and the amount of applause my concepts had garnered as the first thing to do when visiting the site. I often did this repeatedly during the day. It was uplifting when a comment was applauded for the first time, and a new comment was always exciting. Types of feedback found on OpenIDEO are described in Table 4.1.

Table 4.1: Types of feedback found on OpenIDEO

Category	Description	Example
Official feedback	Feedback from the site or official representatives of the site	Congrats! This post is today’s featured concept!
Encouragement	Complimenting and giving positive feedback on the contributions of other users	Awesome concept! Totally believe in it! Keep it up! :)
Reflection	Reflection on the comment or idea from different viewpoints, discussion about the implications and related stories	I agree with Meredith. In general, I think the best parts of a city should be exploited in a positive way. I know in the case of Spain, many of the younger students receive their education or pick up a trade, and go to a different country for work. I think it would be key to involve locals, especially the younger generation, and incorporate activities/events they want in a city.
Questions	Asking questions and requesting more information or clarification	Do you have any specific ideas for what might motivate and incentivise manufacturers to design in this way?
Thanks	Thanking other users for something they have done	Thanks to Amanda, Ally, and Sushmita for raising the security angle.
Acknowledgement	Acknowledging the contribution of someone else in discussion or submission	Angeliki’s concept of crowdsourcing for small tasks is also really great, so along with the legal support net the ‘Assist’ service should also incorporate a micro-tasking network.
Rules	Feedback and instructions about the rules of the site and how people should behave	No apologies on a collaborative platform, Paul;^) We’re all learning here – and you’ve certainly taught me a lot!
Tasks	Suggesting tasks in combination with feedback	Twitter based medical care sounds like a great idea! Consider posting it in our Concepting section.
Mission accomplished	Announcing what has been done, often in response to a suggestion to take an action	Nice find (with the timberland green index)! I’ve added a shot of that to this concept.
Applause	Applause given to inspirations, concepts and comments. Similar to the like-button on Facebook	Applaud button, number of applause counts visible at the site
Number of views and comments	Quantitative feedback on attention and interest	Number of views and comments visible at the site.
Evaluation	Challenge specific evaluation of shortlisted concepts	Multiple-choice questionnaires and bar charts.

4.1.3 OpenIDEO innovation process

At OpenIDEO, the innovation process is considered to be a collaborative learning process. Sharing of information and collaboration are encouraged above competition. Each phase has a deadline before which the contributions to that stage have to be made. The innovation process has several well-defined phases. Except for the early variations, the structure of the process has remained stable from challenge to challenge, although depending on the challenge some phases may be left out. Figure 4.3 depicts the OpenIDEO innovation process as it is presented on the website.

Figure 4.3: The OpenIDEO innovation process

In addition to the public phases, the process also contains implicit phases taking place behind the scenes. The full process, including the implicit phases, is described below. Depending on the phase of the challenge, users can submit inspirations (the Inspiration phase), submit concepts (the Concepting phase), update their own, shortlisted concepts (the Refinement phase), and evaluate concepts (the Evaluation phase). Commenting on and applauding inspirations and concepts on blog posts is possible at any time, even after the challenge has finished. Other activities are possible only for a limited time, during the corresponding phase.

Challenge design. As stated earlier, the activities on the site are organised around challenges, which are usually related to large and complex environmental or societal issues. OpenIDEO receives the funding to run the challenges from challenge sponsors, who can be organisations or individuals. The sponsor pays the costs of running the site and the salaries of the facilitators. Although I have not found an explicit description, I suppose that the challenge is designed in collaboration with representatives of the sponsor and employees of OpenIDEO. The resulting challenge brief is posted on the site, which marks the beginning of the challenge:

> Organizations or individuals can sponsor a design challenge, as long as it's for social or environmental good. All OpenIDEO challenges require financial sponsorship to help underwrite our own costs associated with managing the challenge and providing tech and community support (OpenIDEO 2012b).

Table 4.2: Description of the OpenIDEO innovation process

Phase	Description	Output
Challenge design	Before launching a new challenge it is designed in collaboration with representatives of the sponsor and employees of OpenIDEO.	Challenge brief
Challenge brief	The challenge brief describes the design problem, context and goals, and marks the beginning of the challenge. It is usually a combination of a written description and a short video featuring a representative of the challenge sponsor.	
Inspiration	The Inspiration phase consists of two related tasks: learning as much as possible about the problem and finding examples of solutions that have worked elsewhere.	Inspirations
Synthesis meeting	After the Inspiration phase, the OpenIDEO team, possibly with the help of representatives of the sponsor, hold a synthesis meeting, where they group inspirations into emerging themes.	Themes
Concepting	New ideas are generated for solutions to the problem described in the challenge brief.	Concepts
Applause	The users are asked to help select the best concepts for further refinement by applauding and commenting on the concepts they like.	Concepts ranked by views, comments and applause
Shortlist selection	The OpenIDEO team first reads through all the concepts and comments and takes note of the applause given for the concepts. Usually, the team then selects up to 20 concepts for refinement.	20 shortlisted concepts
Refinement	The shortlisted concepts are improved in a collaborative fashion.	Refined concepts
Evaluation	Users evaluate all the shortlisted concepts against specifically developed evaluation criteria.	Evaluated concepts
Winner selection	The OpenIDEO team in collaboration with representatives of the sponsor decides who are the winners.	Winning concepts, sometimes challenge reports
Winning concepts	The winning concepts are announced on the site.	
Realisation	The Realisation phase is about telling stories and the dissemination of information about implementations taking place outside the site. Implementation of developed concepts is outside the scope of the platform.	Reports on implementation

Challenge brief. The challenge brief describes the design task at hand and marks the beginning of the challenge. It usually combines a brief written description and a short video featuring a representative of the challenge sponsor. The challenge brief describes the problem and context, explains why the issue is important and states the general goal of the challenge. The sponsor of the challenge is also presented. OpenIDEO typically features complex, ill-defined problems, for which it is unlikely that a single solution exists. The challenge brief does not give exact criteria for a successful solution. An important part of the design task is to figure out what exactly is the problem.

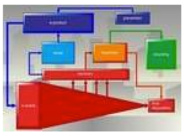
Inspiration. This phase consists of two related tasks: learning as much as possible about the problem and finding examples of solutions that have worked elsewhere. The challenge brief only gives a short introduction to the issue and it is the task of participants to figure out what exactly should be solved and how. During the Inspiration phase, the problem is defined in multiple ways in inspirations posted to the site and in comments on these inspirations and the challenge brief. This phase is, in essence, about preparation for idea generation following the Concepting phase. The main task of the participants is to use a specific Inspiration form to post anything interesting, as long as the content is related to the challenge brief. Inspirations may contain definitions of the problem, such as short videos describing some aspect of the issue, personal experiences, or existing solutions to a similar issue. In addition to common tasks, each challenge usually features special missions. These are challenge-specific tasks to be performed during the Inspiration phase. They give participants more detailed instructions on what kind of activities to take on during the Inspiration phase. Missions are given rather indirectly in the side bar of the site and users can participate in them as they wish. Inspirations can be tagged to missions, which seems to be the only way the completion of missions is tracked. Often, missions relate to finding out more about an issue, revealing personal experiences and stories, and finding working solutions to similar problems. The most frequently-recurring mission is called Surprise us, for which OpenIDEO asks users to submit almost anything interesting, even if it is not that closely related to the challenge.

Synthesis meeting. After the Inspiration phase, the OpenIDEO team, possibly with the help of representatives of the sponsor holds a synthesis meeting, where they group the inspirations into emerging themes. The themes and accompanying descriptions are then posted to the website at the start of the Concepting phase. The team may also define a set of design constraints or other further instructions for this phase. Since the sixth challenge (How might we improve maternal health with mobile technologies for low-income countries?), the Concepting phase has been accompanied with a challenge-specific Brainstorm-in-a-Box toolkit.

Concepting. The main goal of the Concepting phase is to generate new ideas for solutions to the issue that is the topic of the challenge. Usually, the facilitators combine inspirations with challenge themes, which are supposed to give directions for conceptualising. Usually, there are up to half a dozen themes. ‘How might we … ?’ questions are sometimes used, as shown in Figure 4.4.

The concept themes give direction to conceptualisation. Content and discussions from the Inspiration phase are used to point at possible solutions. Users are asked to take what has been learned from the Inspiration phase and use it to generate new ideas, which are then posted to the site using a specific Concept form. The form provides instructions to the user on what to include with the concept. The empty fields, questions and possibilities for uploading content all provide hints on what is expected from a good concept. A visual approach and use of pictures is heavily encouraged; the site facilitators regularly post comments encouraging users to add visuals to concepts without pictures or videos.

TRACK THE LIFE CYCLE



When we buy an electronic item, we don't always consider what went into making it or where it goes after it leaves our hands. Tracking and certification – through data collection, mapping and more – can help us better understand and trust the supply chain that makes our electronics and that disassembles them.

- How can we use technology to certify and verify the life cycle of our e-waste?
- How might we use data or mapping to better connect consumers with the story of their electronics?

Figure 4.4: Example of a theme from the 'How can we manage e-waste & discarded electronics to safeguard human health & protect our environment' challenge

Applause. During the Applause phase, users are asked to help OpenIDEO employees to select concepts for further refinement by applauding and commenting on the concepts they like. According to a forum post by one participant, applause is about appreciation rather than being a popularity poll. In addition to the official purpose of supporting decision making, applause is used to give feedback to participants for their efforts. This is possibly the most important reason for this phase. I personally found applause to be an important form of feedback. After submitting my first concept, checking whether my concepts had gained any applause was the first thing I did during subsequent visits to the site. The Applause phase normally lasts for a week, but the deadline is often extended.

Shortlist selection. Before the Refinement phase, the OpenIDEO facilitators read through all the concepts and comments and take note of the applause given for the concepts. Although the number of comments and the amount of applause are used to help shortlist concepts, the facilitators still read through everything to spot any hidden gems. Typical challenges include between 100 and 300 concepts, with 600 being the maximum number of concepts submitted to a single challenge. Examining all of them requires considerable effort but presumably it is not overwhelming, especially if several facilitators can be used. According to my best estimate, there are currently seven people working at least part-time as OpenIDEO facilitators. Representatives of the sponsor participate in decision making and normally up to 20 concepts are selected for the shortlist. The owners of these concepts are asked to refine them in the next phase. The general impression is that the concepts are selected based on the merit the decision makers see in them.

Refinement. In the Refinement phase, shortlisted concepts are improved in a collaborative fashion. Users can update their concept based on feedback from facilitators and comments they have received from other users. Users may develop rough prototypes of their concepts, such as simple websites and mock-ups. Some of the users go much further still, especially in cases where they are actually planning to implement the concept on their own. Comments and discussion between participants play an important role in the development of concepts.

Evaluation. In the Evaluation phase, users are asked to evaluate all the shortlisted concepts against specifically developed evaluation criteria. Facilitators define the evaluation criteria before the Evaluation phase, possibly with the help of the challenge sponsor. As with other instructions, the evaluation criteria are defined behind the scenes by OpenIDEO facilitators, possibly in collaboration with representatives of the sponsor and then posted to the site at the beginning of the Evaluation phase. A specific Evaluation form is used for evaluation, which usually features three to five questions and a numerical evaluation for each item with multiple criteria. An example of a few evaluation questions is shown in Figure 4.5.

Figure 4.5: Evaluation of concepts at OpenIDEO

Winning designs. OpenIDEO facilitators in collaboration with representatives of the sponsor decide who are the winners. User evaluation supports the decision making but does not dictate the results. Selecting winners is repeatedly described as a difficult task and it does not strictly follow the order of evaluations. Depending on the challenge, a report may also be prepared; in one anecdotal case, IDEO used the platform to support its own project work. Results from the challenge are included in the project report. Winning concepts are then announced on the site. Apart from the glory and congratulations from the community, the winners are not rewarded, although sometimes regular winners and active participants are featured in a short interview on the site blog.

Realisation. The last phase of the OpenIDEO innovation process is realisation. The Realisation phase is about telling stories and disseminating information about implementations taking place outside the site. Users are asked to submit stories about successfully implemented concepts. The site works as a platform to disseminate information about concepts that are going forward. Implementation of developed concepts is outside the scope of the platform. OpenIDEO does not implement anything on its own. Challenge sponsors may implement some of the concepts, but much information is not available. Users of the site are encouraged to try implementing concepts by themselves. Some of them do indeed build prototypes and test and develop them in their own communities.

4.1.4 User experience at OpenIDEO

At the time of visiting, OpenIDEO had two challenges available, both near the end of the Concepting phase. Clicking the challenge link anywhere on the site takes the user to the current phase of the challenge. The whole process is depicted visually on a prominent place on the site. It is clear, bright, and immediately visible, giving a lot of information in an easy to access format. Each phase is described below the process chart in a concise and easily understandable way. The whole site is very visual and the use of images is encouraged throughout. The site has a lively feel, with many images and bright colours. At first it feels big and complex, but the size of the site soon turns out to be manageable. It is possible to scroll through all the submissions in a given challenge.

It seems that most users, or at least the prominent ones and the ones that are promoted on the site, are highly educated professionals and designers. Many of them work on jobs related to business development or design, or in an area related to specific challenge topics, such as mobile application development in rural India. The users of the site are definitely not a random sample of Internet users. Instead, self-selection favouring design professionals, higher education and social innovation seem to be at play. For many participants, the OpenIDEO community provides a venue to collaborate with a global community, which seems to be one of the reasons why people participate on the site. The social ingredient of meeting new people offers users opportunities for bouncing ideas back and forth, for learning from others and relating the ideas of others to their own projects.

Participating in OpenIDEO requires considerable motivation and effort from the user. In the beginning, the site can feel overwhelming, and it is difficult to know where to start. Due to the nature of the issues discussed on the site, it requires effort from the start. The user has to develop an understanding of the issue at hand before making any meaningful contributions, apart from applauding concepts and other content. In the challenges I participated in, I was left with the feeling that I was only scratching the surface. To gain any real insight, I would have had to work much, much harder. The same thing happened after the Inspiration phase, where users are supposed to come up with new concepts. The user has to figure out on his or her own how to best use the contributions of other users. These feelings were also reflected by other participants of the site: one mentioned it that it was overwhelming to see so many creative inspirations and concepts posted. Another user was sure that everyone found it difficult to keep up with the hundred concepts in that particular challenge. Developing concepts, although satisfying, also felt like a lot of work. Some participants even mentioned scheduling freelance work for the Concepting phase.

Refining the concept during the Applause phase was easier; here, it was possible to leverage the community. I received a couple of comments and then asked for suggestions on how to develop the concept. This resulted in two detailed replies pointing to many related concepts that could be used to improve my idea. Here, OpenIDEO reduced the amount of work I had to do: instead of going through all the concepts by myself, I could rely on the community to search relevant inspirations and concepts for me.

Creating inspirations. I participated in the Inspiration phase in the 'How might we identify and celebrate businesses that innovate for world benefit – and inspire other companies to do the same?' challenge. The site already had 275 inspirations (the total for this challenge was 456 in the end), which was a little alarming. I read through the guiding principles, which appeared to be challenge-specific. In my experience, the user needs to understand a lot to participate in OpenIDEO. There are few cognitive shortcuts for the user, which would allow him or her to understand the issue at hand faster and more efficiently than when working on it alone. The collection of information is probably more effective, but every user still has to go through a large part of it to gain a sufficient understanding of the challenge. I browsed through all the inspirations submitted so far in list view to get an overview of the big picture, hoping my subconscious mind would pick up some interesting patterns. The next day I browsed through all the notes related to sustainability, business and environment on my Evernote. I found two interesting notes, which I turned into inspirations for the challenge: sustainability in NASCAR and a special report by MIT on sustainable business. In addition, something reminded me of a TED talk by Derek Shivers, which I added as an inspiration. Creating inspirations was fun and easy, but quite time consuming. I used copy and paste, added links to the original source and decorated the submission with images from Flickr or my own albums. The end results looked pretty.

Applauding inspirations. In addition to submitting inspirations, I took on the task of applauding inspirations submitted by other users. It was not requested anywhere but for some reason I felt I should do it. It turned out to be unfeasible to browse through all the inspirations. Just browsing in list view takes effort and quickly gets boring. Concentration and interest in the task are lost. I found it much more useful to use the default filter, Fresh and surprising, which randomises the order in which the inspirations are presented. I applauded the inspirations I found interesting, most of the time just based on the image, headline and short summary. I browsed through about 160 inspirations, at which point I lost interest. I barely looked at the images. In my experience, I felt I was doing a poor job at applauding. It felt as if applauding did not really matter, and no matter how much I applauded, there would still be a huge number of inspirations left.

Understanding the challenge. I joined the OpenIDEO site when two challenges were in the Concepting phase. The first task to do at this point was to find out what the problem was. The best way for me to do this was by watching videos, such as documentaries on e-waste and TED talks. These gave me details and deeper understanding. Another way was by browsing the inspirations in list view to get the big picture of the issue at hand. Before creating solutions, it is necessary to understand the underlying patterns. The only way to do this on OpenIDEO is to personally study the problem. The site offers a large collection of relevant material but, apart from grouping the inspirations into challenge themes and loose prioritisation through applause and view counts, it does not improve learning much compared with surfing the web alone. I considered using sticky notes to collect the inspirations and organise them in patterns, but never did, because it felt like too much work. Overall, the user is left on his or her own in working out how to best use the collected inspirations. Should I just skim the list or read a random selection thoroughly? The quality of inspirations varies. Some are poorly formulated and uninteresting, whereas others show creativity by taking ideas from other contexts and applying them to the challenge at hand. The site offers many filter options: recent, applause, views, comments, missions and themes. Nevertheless, I was left wondering which was best to use. It felt as if whatever I chose, meant that I was going to miss a good inspiration.

Creating a concept. During the Concepting phase, I wanted to create a concept, but did not know where to start. That was partly because I joined in the middle of the process. I thought that reading the inspirations might be a good way to get accustomed to the problem and find ideas for the concept. I tried this approach on a couple of occasions. At times it felt frustrating because of information overload: there were just too many inspirations and I instantly felt like giving up. I felt that I did not have the necessary insight on the problem to make a contribution. Later, I decided to browse through the most applauded concepts and applaud and bookmark all the interesting ones. This way I could gather pieces for a concept. I used Evernote to collect ideas for the concept. I tried to come up with a concept, but it would have needed additional effort and I felt tired, so I decided to do it the next day. I was totally out of ideas for a concept. A few days later I woke up with an idea for the concept. After that, developing the concept was actually quite enjoyable. At this point I found it important to ensure that no-one had scooped my idea already. I browsed through all the concepts in the challenge and winning concepts from the previous challenge, which had similar topics. After seeing a reminder, I checked the challenge brief again and looked at what exactly was asked in the submission form. I wrote a description of the concept and created a very simple prototype website. I built my concept, making use of many inspirations and concepts, using the linking features of the site. After honing the text and adding an image, I submitted the concept to the challenge just in time before the deadline. I felt satisfied with myself.

Applauding. Applauding a concept or inspiration was one of the first actions I carried out on the site. The first action makes the terms and conditions for the challenge (different from the general terms and conditions of the site) pop up and they must be accepted before continuing. I accepted them without reading. In the official Applause phase I had to force myself to applaud anything. Another participant also mentioned in the comment section of a concept that many people disappear at applause time, and that he too had to motivate himself to do it. I did most of the applauding during one visit. I browsed through all the concepts, opened the interesting ones and applauded the ones I liked. I found it hard not to look at the amount of applause the concepts had already gathered. I felt that seeing the current level of applause had an impact on my decisions. I noticed that I applauded concepts both for right and wrong reasons. The right reasons included cases where effort had been put into the concept or a nice video. This was based on feeling, details from the headline and summary, skimming through the concept and liking it, and appreciating the effort. On the other hand, wrong reasons included not applauding a concept because it already had plenty of applause and I was jealous, or out of pity when a concept did not have any applause or comments, even when I did not like it.

During the Applause phase, I also had an opportunity to refine the concept I had submitted to the e-waste challenge. A facilitator commented on my concept and said she liked how it could be combined with many other ideas. I thanked her for the compliment and asked if anyone had any specific suggestions as to which ideas my concept could be combined with. Another user wrote two comments suggesting many possible concepts that it could be combined with. Reading the concepts and accompanying discussions was hard work. I then used sticky notes to work out how concepts could be combined. The updated concept was quite good, in my opinion. I also replied to a few other comments on my concept. This was the most active discussion I had on the site. In the end, my concept did not make it to the shortlist.

Refinement. I still feel that participating at OpenIDEO is hard work, especially in the Refinement phase. Just reading through the potential concepts takes time. It feels as if everyone has to do a lot of work to get an understanding of the whole issue, and the concept they are developing in order to contribute meaningfully. I browsed through all the concepts that made it to the shortlist, which got boring in the end. As a result, I was not very thorough all the way through. Then I tried to figure out which concepts I might have something to contribute to. I noticed that OpenIDEO provides very good feedback for shortlisted concepts on how to develop the concept further. Specific questions to be answered and links to refinement instructions direct the development. It seems as if there is not much activity on the site during the refinement phase. For example, in one concept the latest comment was from OpenIDEO announcing that the concept had been selected for the shortlist. At that time, there were only three days of the Refinement phase left. According to the activity feed, the tenth last activity took place the day before. All this made me wonder how active the OpenIDEO site is. It looked like very little had happened since the announcement of the shortlist. Is this typical? A contrasting viewpoint can be found in an older blog post, where a featured user describes the Refinement phase as being particularly intense.

According to the instructions, this phase is very much concerned with visualisation, clarification, mock-ups and prototypes. I got very frustrated with the platform. To contribute something, I would have needed to read through many long concepts and discussions to find those that would match my skills and knowledge. Then I would need to take up the serious challenge of creating a business model or something else challenging and time consuming. Creating a prototype is asked for in the Verified Skills Academy. I considered this to be barely within the limits of my skills. The golden e-waste concept was very long and as a next step it would require a more honed business model and a pitch to investors, which seemed too much to ask. The Regenerate concept would also need refinements in the business model, which felt very difficult. In Digital data transfer and elimination services, a survey on user perceptions was requested. A few other concepts also lacked a business model. It felt as if participating seriously in the Refinement phase would require real commitment. I would only put in that much effort if I were planning to develop a concept for my own business. This makes me think that perhaps OpenIDEO actually works by connecting the people who are already working on these concepts and brings them together on a common platform. In the end, I did not really participate in the Refinement phase, even though I was supposed to. There was not much going on at the site, at least around the concepts I was interested in. The required tasks were dauntingly large and challenging considering the amount of time I had. Nevertheless, I do wonder whether I am somehow biased in thinking that participation in OpenIDEO is hard work.

Evaluation. The final stage was to evaluate concepts. I could also still comment on or applaud them, but I did not feel like doing so. When I was first exploring the site, I came across a user stating that before the Evaluation phase one should first read through all the concepts and then think about them. I did not want to do that. I simply wanted to get something done quickly. I decided to start evaluating the concepts in order and to continue as long as I felt like it. The evaluation questions were confusing. The results were available only after I had finished the evaluation.

I then decided to evaluate the concepts that had the least number of evaluations. Most of the time I skimmed the concepts and evaluated them based on their top level descriptions. Again, my evaluations were inconsistent and subjective. I found it difficult to be critical. Even if I did not like the concept I often gave it a mid-range evaluation just to cheer up the user submitting the concept. I gave one very positive evaluation partly because the concept was similar to the concept of my own that did not make it to the shortlist. If a concept was lengthy, I only skimmed through it. I skipped evaluating one particularly long concept completely. It would have been unfair to evaluate it without even reading it. In general, I did not put much effort into the evaluations. Making a choice between the options was not difficult as there were only three options. The wording of the choices made it easy to choose. I got the feeling that problem definition was the difficult part in OpenIDEO challenges. It felt as if people did not get it right, because there were many interacting issues. I have seen several concepts on the shortlist that I consider doomed to failure. They simply do not make sense when taking into account company policies and interest groups, not to mention the logistics issues required to upgrade electronics en masse.

4.2 Quirky

Quirky describes itself as follows:

> We make invention accessible. We believe the best ideas in the world aren't actually in the world... they're locked inside people's heads. We exist to solve that problem. (Quirky 2013c)

Figure 4.6: Quirky Live Evaluation (Quirky 2013a)

Quirky is devoted to making invention accessible by allowing inventors to submit their ideas to the Quirky website. The company then selects a few ideas each week for further development and commercialisation by Quirky’s own professional designers and engineers. Initially focused mostly on relatively simple consumer products, the company has lately expanded to include products involving electronics. Examples of product categories include kitchenware and smartphone accessories. The professional designers and engineers employed by Quirky do the most of the development work apart from coming up with the original idea. This is different from OpenIDEO, where employees work mainly as facilitators, leaving participating users to develop the concepts. The majority of tasks that users can carry out at Quirky are well-defined and relatively simple. Users earn influence points from each task they do. If a product is commercialised, the influence points are turned into a share of profits from the sales.

4.2.1 Rules at Quirky

Joining the Quirky community is free. Quirky's mission is to 'Make invention accessible'. Their 'job is to act as sort of shepherds of our inventions' (Quirky 2013c). For Quirky to go forward with an idea, the community has to like it and the idea also has to solve a problem. Once accepted, the product license and copyright become the property of Quirky, as explained in the very lengthy set of terms and conditions. Quirky does not patent its ideas, but patents may be used as inspiration for ideas, especially those from its partner General Electric (GE). The company runs a regular weekly schedule – 'This invention machine stops for no one, so our live product brainstorm will still be held at its regular hour (6 pm)' (Quirky 2013d) – with events related to product development. The website follows a consistent development process. Although the company emphasises collaboration with users, the vast majority of work appears to take place behind the scenes, carried out by the company's employees (professional designers and engineers). Most tasks that users are asked to carry out are rather simple and quick. Timeframes vary: idea evaluation is open for 30 days for each idea, whereas research, naming and design tasks are often open for only 2 to 7 days, or even less. The time available for completing the tasks is shown with counters. There may be limits to the number of contributions per user for each task. Examples include 15 votes per day for idea evaluation, 3 votes for design projects, or 1 submission to a design project. In general, the participants of Quirky (Quirky 2013e) are expected to:

1. Stay active.
2. Participate in everything: Influence from both the smallest actions and largest contributions.
3. Converse in the forum: Remember to be civil and open minded!
4. Message other members: We encourage communicating with members about their ideas etc. but avoid filling inboxes with promotion for your idea. (User needs to follow you before you can send him a message)
5. Follow and be followed.
6. Give props where props are due: When community members do a good work, give 'em a pat on the back
7. Specify your skills: Now's not the time to be humble, show off your skills!
8. Use your real name: We offer alias function, but we prefer real names
9. Think outside the box: Quirky sales = money in your pocket
10. The more you hype up Quirky and get people to click, the more influence / cash you'll earn
11. Promoting this link will earn you credit for actions taken by visitors: Up to 10% for sales, 20% commission on paid idea submissions
12. Find your social sales link: Social sales give you opportunity to stimulate final phase by cold, hard sales
13. Engage with the community: Check back often to see how people are commenting
14. Learn from the best
15. Do your research

Quirky's specialty is measuring influence to define the rewards for participants. In Quirky, influence is a real-time measure of a user's contribution to a project and is used to measure the percentage of the community's total share in a product. Influence generates money: 70% of Quirky sales goes to the company, with the remaining 30% (10% for indirect sales) divided among community members who influenced the product. The inventor (ideator) takes the lion's share, whereas smaller tasks earn less. The breakdown of influence among different tasks at Quirky is described in Table 4.3. There is no hard-and-fast rule or guarantee for earning influence: it is awarded solely at the discretion of Quirky and depends on many factors.

Table 4.3: Rules on gaining influence at Quirky (Quirky 2013e)

Task	**Description**	**Influence %**
Idea submission	If your idea takes off and is picked for development, you'll receive the largest percentage of the community pot when your Quirky product starts flying off the shelves!	42%
Collaboration	Submit a revision using Quirky's collaboration tool: if the ideator accepts your revision, you could earn yourself a portion of the collaboration influence. If no major revisions are submitted and accepted, this percentage is added to the Sales influence total, described below.	6%
Comparison	Product development is competitive, so Quirky encourages users to browse the web for products that are similar to submissions, and to submit them as a Product Comparison. Each submission can have a total of 5, and the 3% influence is split between the 5 contributors	3%
Winning votes	Have you stumbled upon an idea submission that you want in your hands, RIGHT NOW? Vote for it! If the idea you voted for is selected, you'll have earned yourself a piece of the revenue pie. The 6% influence is divided amongst all voters for a winning idea.	6%
Losing vote	The idea you voted on was placed Under Consideration, but didn't make it past staff evaluation. That's okay, we still want to give you credit for pitching in: while the idea you voted for didn't make it, we'll award you influence in every other product that won that week. 3% of each product's influence goes towards this reward.	3%
Research	Quirky conducts market research in the form of surveys: help us out with some answers, and you'll earn a cool percentage of the product's retail revenue. We split the project's 5% influence between everyone that participates!	5%
Design	Our design team wouldn't be able to launch two ideas a week without your help. If you submit an idea for a new feature or improvement that they use in a final design, you'll receive up to 3.75% of that product's influence. Voting for a selected idea earns you influence as well. Winning ideators split 3.75% of the project's influence, while winning voters split a total of 1.25% influence.	5%
Refine	There's always room for improvement, don't you think? We'll open up this phase when we need your help smoothing	5%

	out the kinks in the design process. 5% of the product's influence will be split between those who participated.	
CMF	Color. Material. Finish. Our design team provides the choices, and you pick them! The project’s 5% influence is divided among the members who voted for the winning option.	5%
Naming	Members can submit and vote on potential product names, and the most popular will be made official! The winning name will receive 3.75% of the influence, while winning voters will split the remaining 1.25%. Note! This has changed after the introduction of the Naming Game!	5%
Tagline	A product's name may give it an identity, but the tagline gives it character. Much like the Naming phase, this project allows members to submit and vote on ideas for product taglines. 3.75% influence goes to the winning submission, while 1.25% goes to the folks who voted for it. Note! This has changed after the introduction of the Naming Game!	5%
Pricing	When a new product is launched on the Upcoming Page, a seven day countdown is started. Share your input on the product's price before the timer runs out, and you'll be awarded a portion of influence! The total 10% influence is split between those who shared data within the first 7 days.	10%
Sales/ SKU Selecting	When we launch a product, anyone who contributed to its development will have the chance to earn more influence by being one of the first to buy. Simply purchase the product from the Quirky store within its first two weeks, and you'll add another chunk to your total influence. The amount of influence split between buyers can vary from 15−25%, depending on the number of other phases included in the product's development. The number can go up if we skip a refine phase, down if we add a second design phase, etc.	15−25%

4.2.2 Feedback at Quirky

Like OpenIDEO, Quirky provides many forms of feedback to users. All ideas receive feedback from the community in the form of votes and comments. Ideas that are put Under Consideration (UC) also receive feedback directly from Quirky staff. In an ideal situation, every idea would get good feedback, which is something at least some of the staff would like to fix. However, feedback from staff is not the point of the Quirky process, and ideally there would be no need for it at all. Automatically generated feedback reports are mentioned a few times here and there, but I have never seen one. Most ideas never make it out of community curation. As an example, someone put forward the idea for an edible Frisbee, but this received the following feedback: ‘Fortunately we had 12000 community members tell the inventor his idea was the stupidest thing they ever heard’ (Fenn 2011). The comments on submitted ideas contain lots of feedback from the community. Many users appear to make positive comments almost automatically on all the ideas they view. Announcements that a user has voted for an idea are very common.

'Voted!' and 'Voted! Please check my idea!' are typical comments. Although, these comments have a positive tone, behind a thin veil there is an implicit request for reciprocity.

The ideas in evaluation can be sorted by the Most Active, which can create a feedback loop of votes. A few users complain about this on the discussion forum (Quirky 2013f): 'Once you get on the most active roller coaster you keep riding' or 'The most active section must be fixed. Same 10 ideas stuck on top', for example. I was also susceptible to this feedback loop on a couple of occasions. I voted for an idea because it had 92% new votes and the idea seemed to be generating a lot of activity. Another time I voted for an idea only because it had over 400 votes and I just wanted to be part of it if the idea was accepted. The feeling is that so many people cannot be wrong. In the Naming game, the system shows the current ranks of evaluated names after the click. Unlike the Most Active category for ideas, feedback in the Naming phase does not give away all the results, as the user cannot see the ranks of all the names. The Pricing game shows statistics on the current projected price and gives a comparison of aggregated and user choices in a bar chart, which creates a feeling of getting the price right or wrong depending on the results, and can lead to adjustments in pricing tasks. It is possible to see how each product is selling in real time. A special type of feedback on Quirky is the influence scores that the website gives to users for participating in various tasks. The rules on how and when influence scores are earned are quite specific. Users are aware of the scores and sometimes even complain about situations where it appears that Quirky is getting something useful from the users (opinions, comments) without giving out influence scores. In the end, if the product under development is commercialised, the influence scores are converted to a share of profits from that particular product's sales. Finally, the disappearance of task items can also be considered as feedback. There is certain satisfaction about a task disappearing from the list, or an evaluation task vanishing after the click of a button.

Sometimes, members of the Quirky staff provide feedback to the community through streamed shows, such as Feedback Friday and Q&A sessions. The Feedback Friday I followed focused on 'the wide world of cleaning products' (Quirky 2013g). The main message was that Quirky wants problems that haven't yet been solved. The most important task is to prove a problem exists and that there is no (current) solution. The community tends to focus too much on derivatives of existing ideas and not enough on the problem (Quirky 2013b). In addition it is important to pay attention to technology and carry out a preliminary research on feasibility.

If a product is manufactured and put on market, the ideator gets his photo on the package and influencers have their names mentioned on it. In one case, this meant 300 names on the package. The different types of feedback found on Quirky are described in Table 4.4.

Table 4.4: Identified types of feedback at Quirky.

Category	Description	Examples (taken from the Quirky website)
Statistics	Statistics on user activities, both individual and at the community level.	This Week's Top Earners - email 1 Jake Zien Earned:$26,160.45 2 Lemonheads Earned:$6,198.09…' '68 members received influence earning 60% influence'
Automatic feedback	The site generates automatic feedback in response to user actions.	You've completed enough of this survey to earn full influence!
Disappearance as feedback	Making things go away can feel satisfying.	While checking what pod power is, I did the pricing task just to make it go away. Disappearing items in the Naming game.
Staff feedback	Feedback directly from the Quirky staff. Usually as comments on ideas, blog posts, messages on the discussion forum, or little notes in product timelines.	This is Steve's first invention with Quirky. Anna and I enjoyed Skyping with him. Nice to get to see our international inventors in real time
Good luck	A user wishing another community member luck with his/her idea.	Good luck!
Congratulations	A user or Quirky staff member congratulating a community member for getting an idea to the next stage.	Congrats on UC!
I like this	A user commenting that he or she likes an idea.	I really like this idea, for some reason the robot from *Lost in Space* comes to mind…
Positive	Simple positive comments on an idea. May contain explanation for why the idea is good, but often only states the functionality of the idea, if even that.	Clever. Good! This is a great idea. It takes customizing footwear and accessories to another level. Cool concept…
Negative	Negative comments on an idea. Often involves some level of explanation, for example, why the idea is not going to work.	I think this is limited because it doesn't allow for an adjustable water/rice ratio Please don't assist him in this - it is 100% someone else's idea
Agree	Stating an agreement with another user, usually in the Comments section.	Agreed. Agree with DQ… love the built in flour concept! You are absolutely right! I love rice too…
Colours (?)	Comments on colour preferences, especially in comments on CMF project. Perhaps not a separate category after all.	Nicest colors. Definitely the best combo. Red looks good for the button, but from a safety standpoint red could also be used for the casing around the cutting blade. Voted
Problems/ requirements	Attempts to define problems or requirements in response to an idea.	I figure this is the most likely to get packed up & brought to grocery store… However, excessive torque… ensure the product is light enough Yes, I agree some germs are necessary for our immune system. However, Raul, I was thinking more in terms of 'Who the heck wants to clean mould off tub?'

		I am a little worried about the muffin toppers. Doesn't moisture have to escape while it is cooking?
Already done	Pointing out that the idea has already been implemented elsewhere.	I laughed so much when I saw the image. But then I searched for it and found it on 589 blogs declaring it the best idea.
Suggestions	Suggestions for improvements.	Designers need to mix this with ratcheting for perfect grapnel! Wonderful… I'd love to see staff links on track and have the bumps arise out of the music notation. That would teach kids how to read music Fun. I added a suggestion
About users	Assessments on what the users' opinion on an idea or product would be. May refer to people in general, known users (friends and relatives) or oneself as a user. Indication to buy may be included.	What a wonderful invention! Would be essential to so many people of all ages. This would help my sister if this was on the market to buy! Being a whole leaf tea drinker, this is a great way to make sure you get 2−3 infusions from leaves. Great idea! I'd buy!
Would use	Stating an inclination to use the product if it was available.	I would definitely use this in my projects as I hate using a tape measure or ruler
Voted	Stating that the user has voted for an idea. Perhaps the most common comment on ideas.	Voted! Great, voted!
Please check my idea	Requests to check the idea of the commentator. Often combined with voting. Sometimes obvious spam (the same message from the same user on many ideas.)	Voted! Please check my idea. Hello great idea! Good luck with it! Believe in it. Also consider my ideas. Thank you very much. Voted
Thanks	Thanking other users. Usually in response to a vote or a comment.	Thank you <name>! Thanks for your support

4.2.4 Quirky innovation process

The Quirky innovation process is rather complicated, consisting of many phases that alternate between crowdsourced and in-house tasks. The process is depicted in Figure 4.7. It should be noted that during and after the observation period Quirky was in a period of rapid growth. Many features of the process may have changed since the study was conducted.

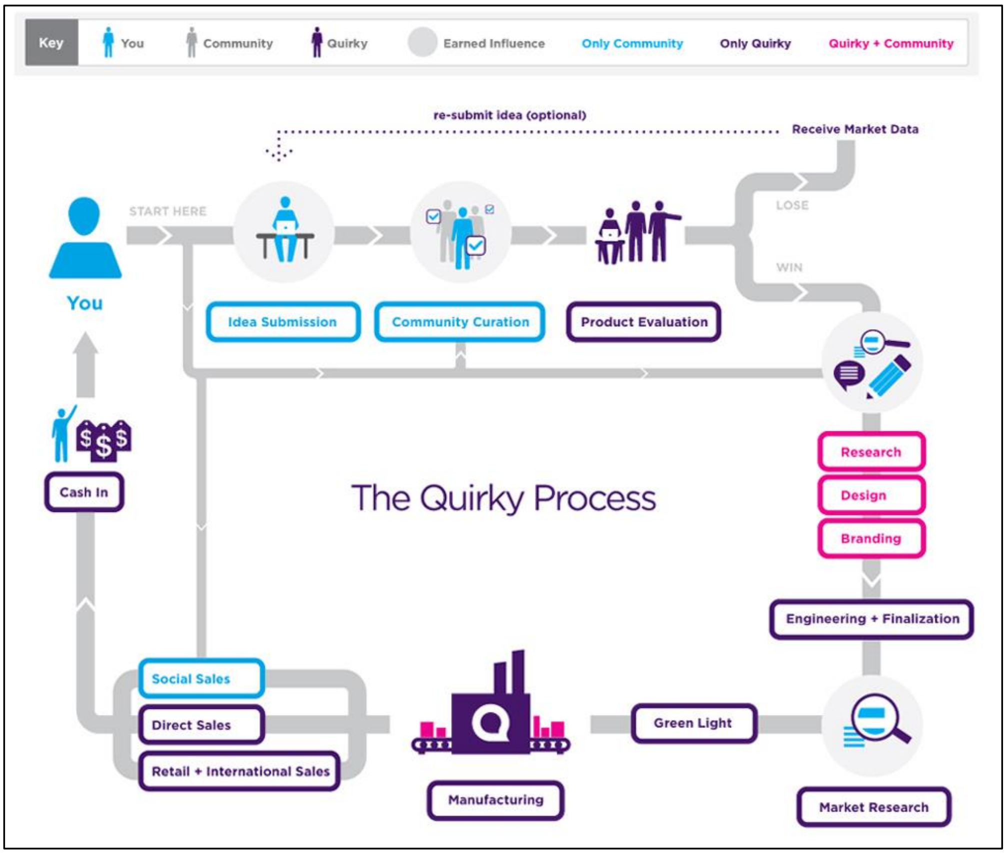


Figure 4.7: Quirky innovation process (Quirky 2013e)

The Quirky website offers many simple and easy tasks for participants. The tasks are mostly well-defined, with varying timeframes for completion. Typical projects that a participant can influence at any one time include: helping choose the next product, helping research a laundry alert product, and helping design a collapsible bike helmet. Nevertheless, a user can sometimes run out of things to do. Apart from the ever-present idea evaluation and the opportunity to vote, the number of tasks available is limited. It appears that Quirky does not need the community to carry out most tasks, but the company has chosen to share these parts of product development with the community.

The vast majority of the work takes place behind the scenes: members of the Quirky staff judge the projects; in-house designers and engineers do most of the prototyping and concept refinement. Table 4.5 gives a summary of innovation process phases at Quirky.

Table 4.5: Description of the Quirky innovation process

Phase	**Description**	**Output**
Idea submission	Quirky community members submit product ideas they have created to the website by filling in a submission form. Photos, drawings and videos can be attached.	Ideas
Community curation	Community views, votes and comments on ideas.	Evaluated ideas
Product evaluation	Quirky staff members evaluate ideas in several steps. A few are then selected each week in an evaluation event that is broadcast live.	Ideas selected for development
Research	Community members fill in surveys created by Quirky staff. The questions often explore potential use cases for the product under development.	Answers to surveys
Design	The community submits and/or evaluates design concepts exploring various aspects of the product under development.	Designs? Evaluated designs?
Branding	Community members submit and evaluate names and taglines for the product by playing a so called Naming game.	Evaluated names and taglines
Engineering + finalisation	Final development of the product. This phase is carried entirely by Quirky employees.	Final product design
Market research	Community members help in defining a suitable price for the product.	Price estimates
Green light	Quirky makes the final decision to commercialise the product.	Decision to manufacture
Manufacturing	Quirky outsources manufacturing.	Products
Sales	Sales of the product through the Quirky website and third parties.	Cash

Idea submission. Idea submission marks the beginning of the product development process at Quirky. Community members submit their product ideas using a form on the website. Submitting an idea costs ten dollars. Sometimes, Quirky features special challenges, which involve a more detailed brief describing the types of ideas that are expected, such as accessories to Apple products. Such challenges may also offer free idea submissions. Examples of past challenges and on-going product categories include: Apple accessories, everyday products, a 24 hour challenge on Apple accessories and ‘Wink: Devices that think for you’. Although not mentioned anywhere on the site, I have a feeling that Quirky is mostly about finding good problems for designers and engineers to solve.

Idea evaluation. The Quirky website provides a comprehensive description of how ideas are selected for development:

> There are two main steps to idea selection: Community Curation and Staff Evaluation. During Community Curation, members of our online community view, vote, and comment on all the Ideas submitted. Quirky Staff then looks through Ideas and, based on many factors including number of community votes and issues raised in community comments, may decide to move some Ideas into 'Under Consideration'. This is the Quirky Staff ʽshort list' of Ideas that we think are promising. The Quirky Community then gets to submit similar existing products, collaborative edit suggestions, and more comments to Ideas in this short list. The Idea is then evaluated by Quirky Staff and thoroughly analyzed in three areas: Design potential, Marketing potential, and Viability. The Ideas that score high in these areas are then put before Quirky Staff in our Weekly Eval meeting where we discuss the Ideas in depth as a group. Some Ideas are chosen to move forward in the Quirky Development Process, some are not. At the end of the day, though, all Ideas submitted will receive feedback from the Quirky Community, and those placed Under Consideration will receive feedback directly from Quirky Staff (Quirky 2013h).

Idea evaluation at Quirky is divided into two phases: community curation and in-house product evaluation. First, the members of the online community view, vote and comment on all ideas submitted. From a user's perspective, idea evaluation consists of browsing submitted ideas, reading idea descriptions and comments, watching possible introduction videos (rare) and commenting and voting. Ideas can be sorted by their newness, amount of activity, evaluation deadline and by their Under Consideration status. The in-house product evaluation starts with Quirky staff picking ideas from community curation and putting them under consideration as a step towards a more thorough evaluation. Every member of staff is responsible for finding the ideas with most potential. A committee then determines which ideas merit further discussion. It convenes for two to three hours each day from Monday to Wednesday to discuss ideas under consideration. Five to ten members of the committee picked from a rolling roster of staff select up to fifteen ideas with the most potential for a weekly live evaluation. The committee looks through the full list of ideas Under Consideration, starting with ideas discussed the previous week or with the highest votes. Each submission is discussed at length from many perspectives. By Wednesday, the Preval committee has a list of strong contenders and maybes, from which it selects concepts to be discussed in the live evaluation. Without a review of each idea, employees risk missing great ideas.

Every Thursday evening, Quirky crowns brand new inventors through their product evaluation process. Ten to fifteen ideas are reviewed live, of which three to five are normally selected for design and probable commercialisation. Most members of the Quirky staff participate along with occasional guests. The event is streamed live over the Internet. The community can participate through live chat or different voting tools. Video recordings of past events are available on the website. During the live evaluation, ideas are presented briefly along with a summary and statistics (votes, etc.) projected on the wall. Usually, the CEO hosts the event and solicits different viewpoints from the audience. After a few minutes of debate the attendees votes by raising hands on the question, 'Should <name> become a Quirky inventor tonight?' Sometimes the community is also encouraged to vote at this point. There are three possible decisions: 1) a design is not built; 2) the idea is explored further or returned to later; and 3) a design is built. If the decision is to explore, the audience gives a single clap. A decision to build something elicits solid applause. Decision making is fast and effective and appears to be reached by consensus. The community participates in the evaluation event through chat, a sentiment meter, thumbs up/down and a pricing tool. The sentiment meter measures how much the community likes or dislikes an idea. Thumbs up/down is used to ask for an opinion on a proposition. The pricing tool is used to evaluate how much the community would pay for the product, which helps to determine whether an idea is good and whether people would be willing to pay for it. After the votes are tallied, they are displayed and discussed by Quirky staff. When Quirky decides to develop a product, employees go through the old idea archives to check whether anyone else has suggested something similar before and so deserves influence.

Brainstorm. Brainstorms are another live weekly event, during which the Quirky design team discusses a newly selected idea and explores its design directions. The community can participate through chat and the event is streamed over the web. Typically, brainstorming results in the creation of a mind map.

Research. In the research phase, the community is usually asked to answer short surveys or questionnaires about the product under development or about the behaviour of potential users. Research tasks are usually simple surveys. One exception was an ethnographic research project, where the participants were asked to recreate a potential use case of a product under development, and to photograph and submit the results to Quirky. Figure 4.8 depicts instructions for a typical research task.

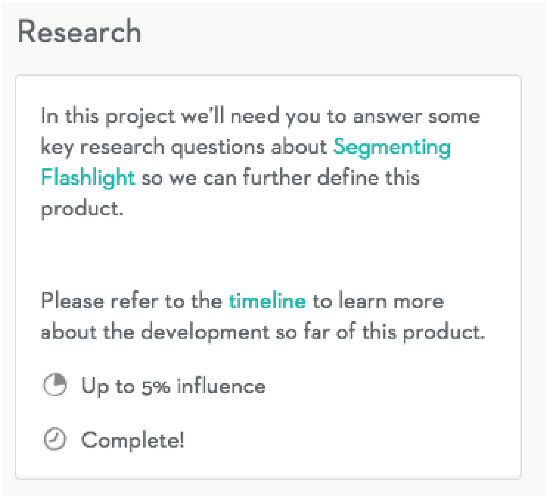


Figure 4.8: Example of research task at Quirky (Quirky 2013).

Community design. In the community design and refinement phases, Quirky is looking for and evaluating concepts that help inform their design direction. Depending on the case, the community may submit design ideas for a particular product or product feature, and vote and comment on it. Organisation of these tasks varies. Sometimes, Quirky has already looked at alternative concepts and the only role for the community is to vote. Sometimes, there is a submission phase for community designs followed by community voting. Sometimes, the submission and voting phases occur concurrently. Figure 4.9 provides an example of a typical community design task at Quirky.

Refinement. In the refinement phase, the community votes on a few design options created by the Quirky staff. It is very similar to the colour, materials and finish (CMF) phase and also shares commonalities with the community design phase. Quirky employees carry out the designs that are evaluated. There are only a few options, in one case only two.

Colour, materials and finish (CMF). The CMF phase is about choosing the aesthetic details for the product under development. Quirky provides a few options: the community gets to vote for its favourites:

> Before we launch, we need your help deciding what color scheme would work the best. CMF: In this project we'll be choosing the colors, materials and finishes for the finalised Rice-For-One. Please refer to timeline for info (Quirky 2013).

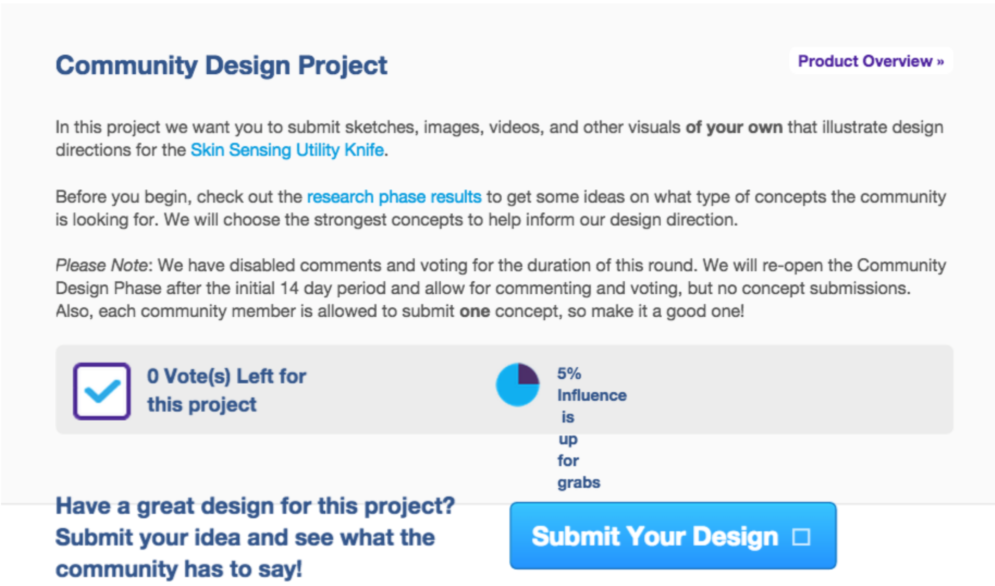


Figure 4.9: Example of a community design project at Quirky (Quirky 2013).

Naming and tagline. Naming and tagline tasks changed during the observation period. At first, they resembled the design task with submissions and voting. However, the process was then changed so that two names and taglines were compared – an approach similar to Wikisurveys at All Our Ideas. This version is based on the ELO rating in chess tournaments. Evaluations compare two options side by side, with one chosen as the winner, as depicted in Figure 4.10.

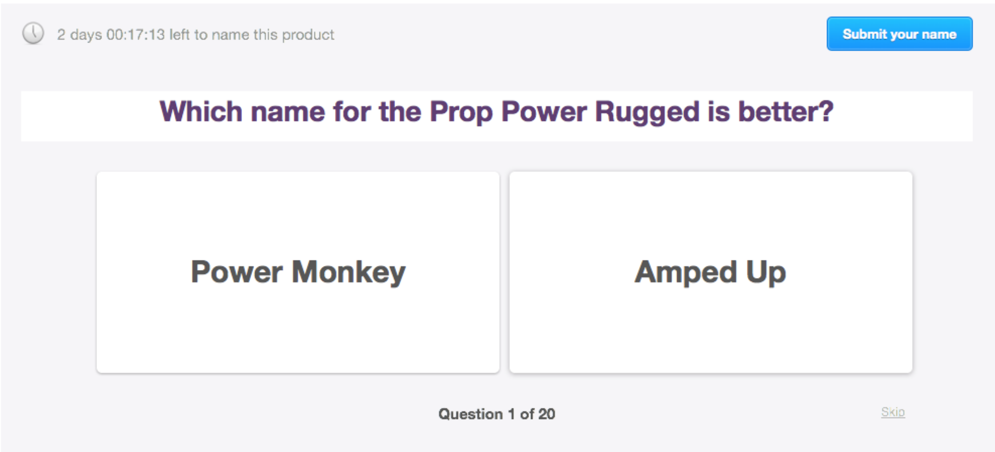


Figure 4.10: Quirky Naming game (Quirky 2013)

Pricing. Quirky uses a pricing game to help determine the appropriate price for new products. Community members are shown a summary of the new product with a picture, and then asked to give their price estimates using four different frames: too cheap, good bargain, a bit pricey, and too expensive. The answers are aggregated across all users and the aggregated results are shown as feedback, with the user's own estimates indicated in the graphs. Figure 4.11 presents the Quirky Pricing game.

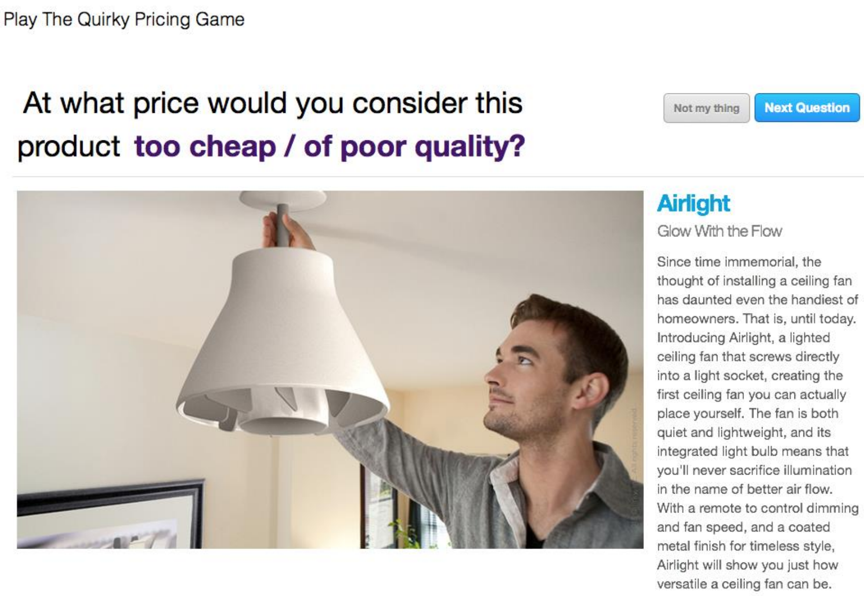


Figure 4.11: Quirky Pricing game (Quirky 2013)

4.2.5 User experience at Quirky

The Quirky website has a friendly style, which is similar to that of OpenIDEO. I found myself being enthusiastic about the site. This could be a great site to learn about product design. The site feels simple and clearly organised. I quickly understood the whole Quirky approach. On first impression, the site felt ingenious. Tasks are well defined, fun and easy, and a new user can quickly participate meaningfully. A possible caveat is that I was already familiar with Quirky. I knew they carried out collaborative product development, where users post and evaluate ideas, with the Quirky team building prototypes and taking care of manufacturing.

It is the user's responsibility to find out what to do next. When checking available tasks on the site, it seems as if there are new things to do every day, in addition to the ever-present idea evaluation. Time frames for tasks vary. It can be difficult to know which tasks have already been completed; the user has to rely on memory. The site has a fast turn-around: new tasks are available almost every time a user logs in.

Quirky neatly aligns the selfish motives of users with its own business goals. Influence is only earned if a user works on something that becomes a product, which gives an incentive for self-selection. I managed to gain some influence on my first visits to the site (0.4% for voting and 0.006% for participating in a survey), which felt good. On the other hand, some users at the discussion forum complain that paid Quirky team members should not expect free input from the community.

Checking the upcoming products section reveals that Quirky makes a lot of products. At the time of the observation, there were 240 ideas in development, although most of them did not have any tasks open for the community. It appears that the vast majority of the work takes place behind the scenes.

Perhaps some of the strongest feelings on the site are generated by different forms of feedback. Getting influence is exhilarating at the beginning. I developed a habit of checking whether anyone had commented on my design contributions. It is easy to understand why people get enthusiastic when their products are selected for development or production.

Idea submission and evaluation. I did not submit ideas to Quirky. Despite trying, I could not come up with anything decent. Instead, a big, friendly button directed me to idea evaluation. The task appears to be about spotting promising ideas, or recognising cool products when they appear. Although there are several filters, the site lacks a filter for most popular ideas. In theory, this should reduce evaluations bias (independence). The number of votes that ideas have gained is not available either. However, I could sort ideas by their activity level, which is measured. Ideas can be browsed in an attractive infinite scroll interface. I always turned this option on. There are plenty of ideas to go through. Ideas can be expanded in place instead of opening them up in a new tab or window, which is useful. Opening ideas in anew window tends to break the flow of evaluation and slow down the evaluation process. According to the forum, participants have developed several approaches or processes for idea evaluation. These include (Quirky 2013):

- When I see an idea I like, I comment. When I want to see ideas I like, I select to see commented. Then I can see all the ideas I like and can change my votes.
- I used to open all the ideas I voted for but lately it's so overwhelming the sheer number of ideas that I only open about half to what I vote for.
- I skim ideas. If I think they have got merit, I will rate them. Then I select options to look at the ones I've rated and choose to vote among those.

Idea evaluation can feel both overwhelming and tedious. According to one participant, the number of submissions had been getting unmanageable lately and as a result, the quality of feedback had lessened. Most ideas submitted to Quirky are of rather poor quality. Finding enough good ideas on which to spend all my 15 daily votes was a challenge. On one occasion, after spending 15 minutes browsing ideas, I had managed to vote for only three ideas, and even those ideas I did not particularly like. It is no wonder that Quirky charges for submissions. At the forum it was claimed that people vote just to earn influence in case an idea goes through. I was also guilty of this behaviour on a few occasions. The Most Active filter helps identify ideas that the staff are likely to select. Forum members see the filter as a good predictor. However, it may be better at measuring the effect of social proof on ideas that the community likes.

Participants in the discussion forum have questioned whether votes have any significance. The most critical participants consider the whole idea of community curation to be fake. They argue that votes completely meaningless and that Quirky selects ideas based on the simplicity of implementation and production. On the one hand, the critics claim that votes do not matter, as ideas with only a few votes can reach live evaluation; on the other hand they claim that the most active ideas with hundreds of votes are selected for evaluation too often. These claims seem to be contradictory. A more balanced opinion is that 'votes are nice, but not as important as working on your project to make it the best it can be' (Quirky 2013). Votes do not count alone: they are simply part of the data. There is probably considerable variation in the number of votes for ideas that reach live evaluation.

People can protest at any divergence from the established process. For example, when a Quirky employee helped a product get to the evaluation phase based on a forum discussion, some members of the community saw this as bad manners. Although disappointments may partly result from not understanding the process, as one participant confesses, the argument about challenges in the evaluation system may still have some merit. From the outside, it seems that considerable effort is taken to separate the wheat from the chaff when selecting promising ideas for development.

Comments and discussions about ideas are usually quite shallow. One member went so far as to suggest that most comments were a mild form of spam: a majority of comments appear to be announcements that a user has voted for an idea, perhaps with some encouragement, and a suggestion that the voter's own ideas should be looked at in return.

People may also be unwilling to post concerns about products to avoid trampling on someone's dream. Getting a product to be chosen for development can be very difficult. Personally, I did not make many comments.

Live evaluation (Live Eval). At the beginning of the live evaluation, a highly enthusiastic CEO explains how it works. After a short introduction of the idea under consideration, a rapid-fire discussion between the experts and audience members (mostly Quirky staff) ensues. The CEO presides. Typical examples (Quirky 2013) include: 'You like it… next!', 'Do you have [an] objection? Speak!' or 'Anyone else [have a] negative opinion?' The CEO often asks for opposing opinions depending on what others think. Often a few comments from the live chat are read out. Sometimes, the community is allowed to vote to register their opinion of the product. The final decision is made by a rough count of raised hands among the audience. After a positive decision, the audience cheers, which I find pleasant. A live chat session with between 40 and 300 participants takes place simultaneously, with the decision-making process streamed live over the Internet. After following it for a while, the chat seems ineffective. There is definitely no time to follow what is going on in the chat area when trying to keep up with the show. The live evaluation is not very popular. The highest number of people I ever saw viewing the stream was about 350. Overall, I liked the decision-making approach with its fast-paced discussion and voting by raised hands.

The weekly brainstorm is another regularly streamed show at Quirky. When watching the show it feels as if the people around the table contribute the most value. The live chat area is just an add-on. Discussions on the chat area are unfocused. I did not want to comment. Research tools and polls during brainstorming sessions would be useful. Currently, the show is just about reading comments from the chat area and seeing designers work. This experience is echoed in one user's criticism on the discussion forums of the fact that the community is only allowed to watch but not participate in brainstorming. The same user claims that the results of surveys are rarely incorporated, that the name and tagline processes are only about marketing and that, as a result, the community has very little effect on the final product.

Research. The research phase mostly revolves around completing questionnaires. Sometimes, it was difficult to know whether I had earned influence by completing a survey. Quirky sometimes incorporates questions from the inventor in their surveys, but these questions tend to be bizarre and probably do not give much useful information. Occasionally, surveys include funny details, questions or answer options, such as an interest scale going from 'snoozefest' to 'peed my pants'. I once answered a questionnaire about the wrong product without noticing.

Community design. In the design phase, Quirky asks the community to submit prototypes, sketches or designs. I participated in one enjoyable design challenge by using Sketchup to make a sketch of a multi-use sensor design with and submitting it to Quirky. I also voted for my own design. Unfortunately, many other users had had the same idea. I voted for them to improve the chances of the cube design and in hopes of gaining influence if one of them was accepted. The design phase of the process seems to create many similar submissions. Quirky's user base appears to be larger than that of OpenIDEO. In one case, there were 770 suggestions for the design of a product. I browsed this mass of ideas in random order and voted for three. Ideas I voted for appeared on top of the page, and I could drop votes if I found something better. This worked as a kind of a search algorithm, which stored the best result found so far. Another time, I went through all the suggestions and voted for three of them. I wasn't particularly interested in any of them, but at least I got something done. The community design phase appears to have some sources of possible bias. Comments on suggestions are visible. On at least one occasion, I nearly cancelled a vote when I found out that somebody had made a negative comment. I often avoided bias myself by consciously avoiding looking at other comments before voting.

CMF phase. The CMF phase is mostly about voting with some commenting on aesthetic options. These tasks are easy. Bias is possible: in one case, I voted for an option just because it happened to have the most comments. Once Quirky announces the winning colour scheme, users may feel that they picked the wrong option if their selection does not match Quirky's decision.

Name and tagline. Helping to choose the name or tagline for products was one of my favourite tasks at Quirky after the introduction of a new system. In the old version of the task, scrolling through names and taglines was overwhelming and tedious. With thousands of name submissions, the process of picking favourites became an ever more daunting task. The new Naming game is much more fun: comparing two options is cognitively easy and the disappearance of items is satisfying. I was delighted to find that it worked in a similar way to wiki surveys, an open source tool for idea evaluation (Salganik and Levy 2012). The Naming game is not as intimidating as trying to look through 30 pages of submitted names. I was unable to see the current rankings of all the names, only the ranks of the names that had just been evaluated. This is feedback, but does not give away the results. Some users like the more immediate gratification. Every time I voted for a tag line with a higher rank, I felt that I had got it right, as if the current ranking was the correct one. Playing the Naming game on the mobile app was even better: it works the same way on mobiles, but has a faster feel and the interface is more responsive. Some users complained about the Naming game on the discussion forum. They do not like the fact that the system does not necessarily assign them the winning submission for evaluation. There is a suspicion that submissions are unfairly rotated. Name pairs and user hates cannot be skipped. They keep coming back in later evaluations. It can be annoying to see the same bad name over and over again. Many of the complaints seem to stem from problems in understanding the evaluation system.

Pricing. The Pricing game feels like a clever solution for assessing suitable prices for new products. It even gives cognitive punishments by asking more specific questions if a user does not give a price but tries to skip the question. The game is enjoyable. It was, perhaps, my second favourite task on Quirky after the Naming game. It is difficult to know an appropriate price. When the system gave me feedback on the prices I felt that I was getting the price right or wrong depending on how close my guess was to the average values. I even tried to adjust my guesses accordingly. After adjusting for differences in currencies (euros vs. dollars), my price estimates were regularly within a dollar of so-called correct value. The game is tempting: after pricing, the only way to move forward is to go to the next price evaluation task. I sometimes did the pricing task just to make it go away.

4.3 Threadless

Threadless is a community that supports making 'great art':

> All designs printed on Threadless are voted on and picked by the community. Users can submit designs to Threadless, which are then voted on for a 7 day period by other users in the community. Once the scoring period has ended, the design receives a score from 1 to 5. This is used as a gauge by Threadless to decide what gets made into a tee! (Threadless 2013)

Threadless is a clothing company with a crowd community, which creates and votes on graphic designs. The company then selects the best designs to be printed on its products and pays a reward to the winning designers. Threadless is best known for t-shirts, but has lately expanded to other product categories. Figure 4.12 depicts the Threadless website.

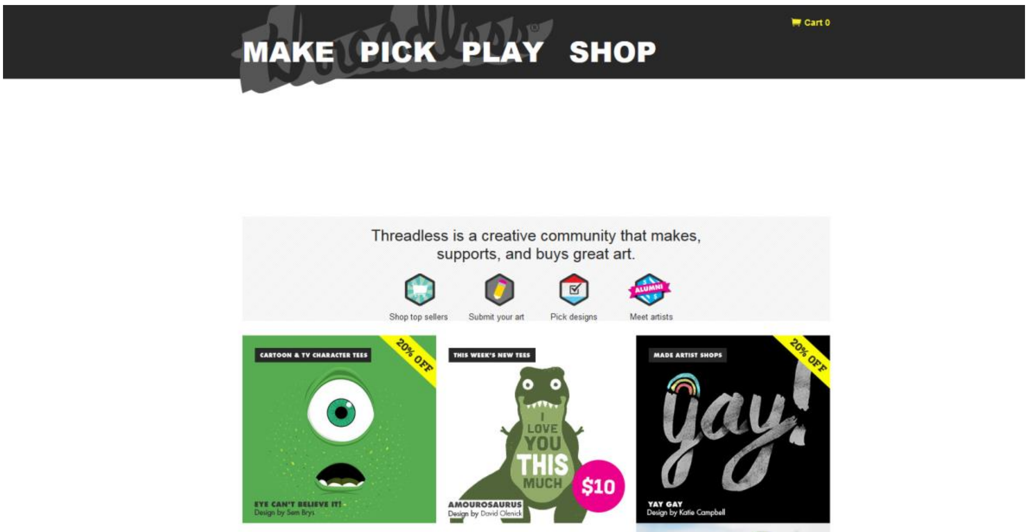


Figure 4.12: The Threadless website

4.3.1 Rules at Threadless

The process at Threadless is summarised below:

> You are Threadless. You make the ideas, pick what we sell, you're why we exist. When you buy from us you support great art. When you join us you support in biggest possible way (Threadless 2013).

Threadless has an optimistic vibe to it, not unlike the one encountered on OpenIDEO. The importance of being pleasant to others on the site is emphasised explicitly, as is the importance of making things. Otherwise, the rules on correct behaviour are implicit and appear to be enforced through the shared culture on the site. This culture is manifested on the Threadless Forum, which is an important part of the site. It offers a platform for the Threadless user community to interact with each other informally. Participating in the community and helping other designers to develop their skills appears to be one of the cultural norms. Participation entails spending time on the Threadless Forum with the community, giving honest feedback on other people's designs and offering good advice, rather than making compliments in return for votes and comments on one's own design. Reciprocation is common, and users often offer to comment on design sketches. The community has a trustworthy image with an implicit ethics code for referencing existing artwork in new designs. Participants do not need to worry about other users stealing their designs. The community also has expectations about the host site's behaviour. For instance, the community should not be exploited by carrying out tasks that could easily be sourced through normal channels. According to the site, no Threadless rules are set in stone. They may be changed based on what actually happens.

4.3.2 Feedback at Threadless

Many of the threads on the Threadless Forum relate to feedback on designs in progress. Feedback from the community appears to be an important part of the design process for many designers: 'Before I start finalising the design, I always get feedback from several people' (presumably including the Threadless Forum) (Threadless 2013). People are expected to be honest with their feedback, but also to show some tact if a designer has put a lot of work into an idea. Although feedback can be harsh, users feels that it is up to the recipient to make constructive use of feedback by doing something with it. The idea is that if you cannot take people being mean or strongly critical of your designs, you are in the wrong field. Opinions appear to be somewhat split on this issue, as other users emphasise the need to be pleasant even to the worst designers. Overall, the main purpose of the feedback appears to be to help people develop as designers. Different types of feedback found on Threadless are listed in

Table 4.6. Most feedback to designers comes through comments from other users on the discussion forum during the graphic design process. Average scores and comments given to designs in evaluation are another important source of feedback for designers. There is even a discussion thread on the forum about the anxiety caused by waiting for scores.

Cheating in evaluation, especially by what is termed downvoting (using multiple accounts to give other designers low scores in the hopes of improving one's own chances of getting printed), is annoying to designers because it ruins the feedback they get from the community about their designs. Compared with OpenIDEO and Quirky, the feedback on designs is more direct and can be harsh, although mostly constructive and in good spirit. Users tend to say upfront if something is not right in a design and usually designers to take the feedback well. The purpose is to help people improve as designers.

Table 4.6: Identified types of feedback at Threadless.

Category	**Description**	**Examples (taken from the Threadless website)**
Positive	Most common type of feedback. Positive encouragement for design submissions or feedback for work in progress. Simple statements acknowledging the quality of designs. This kind of feedback makes the user to want to do more of the same.	This turned out great, will be giving it a $5 when it's up! Wow! Awesome! :) $5!
Negative	Negative statements about designs or behaviour. Aimed at stopping someone doing something.	And god those eyes are terrible If you have to bump your own threads, that's when you know you're onto an idea which sucks No, it's not an oxymoron anyway
Constructive suggestions	Constructive criticism for design and tips and suggestions on how to improve it.	Your second t-shirt mock-up looks way better than the first… Experiment with increasing size on topmost notes, reduce size of flower Try without the white circle, it would looks more natural, I suppose I think rev 2 is ready to submit
Voting as feedback	Scores from voting are visible after the voting period ends. Can be used as a form of feedback, helping designers to gauge how much the community liked the design.	Need votes, need votes, need votes, need votes… I currently had 139 people rate it FYI: mine finished scoring with 2.82
Congratulations	Both official and from other users	Congrats!! You've each won a Threadless pillow! Nice job artists!
Thanks	Acknowledging favours from others	Thanks guys :) Thanks so much for the input, guys! Ok thanks, I'll work on this!
Sales feedback	Feedback and comments on products for sale.	Make a v neck please!!! Amazingly cool idea, and beautiful design! Also as a wall print
Feedback requests	Requests for feedback and help on improving design work in progress	Check out my design I need some tips Please take a look at my design and give me some feedback
Platform feedback	Automatically generated statistics and messages from the platform that give feedback to users on their actions.	You've scored all the designs in the running. Now go find something else to do! I have 52 designs left to score Stats

4.3.3 Threadless innovation process

Out of the three cases, Threadless has the simplest development process. The company describes the process as follows:

1. Get your idea ready to submit
2. Submit to the challenge
3. The community scores your design
4. If your design gets printed you'll get… <prizes>

Users create graphic designs in response to challenges, the most common being the continuously running Threadless challenge, where the design themes are unlimited. During the design process, users may get help from the Threadless community by discussing their work in progress in the Threadless Forum. Once the design is finished, it is submitted to Threadless using a submission form. Certain technical criteria must be followed regarding image quality, size and the use of templates. Threadless staff members review the submitted designs. Accepted designs are then made available for the community to score. Once the seven-day evaluation period is over, the company selects the winning designs to be printed. The process is described in detail in Table 4.7.

Table 4.7: Description of the Threadless innovation process

Phase	**Description**	**Output**
Challenge	The challenge brief describes the theme of designs to be submitted and possible limitations, such as number of colours, only black and white, etc. The continually-running Threadless challenge does not have any specific theme.	Challenge brief
Graphic design	Users carry out graphic design on their own. Typically, this involves coming up with the idea for a design, sketching one or more drafts, line art and colouring. Users often ask for feedback at the Threadless Forum, and refine their design accordingly.	Final design
Submit to challenge	Users submit the final design to a Threadless challenge.	Submitted designs in requested file formats
Scoring	Threadless staff members review the submitted designs. The Threadless community comments and evaluates designs for a week on a scale from 1 to 5.	Evaluated designs
Winners	Threadless staff members select the winning designs to be printed.	Winning designs

Challenge. Threadless hosts a continuously-running Threadless challenge and themed limited-time challenges to inspire artists to create designs they might not have thought of otherwise. External partners may join the themed challenges and offer additional prizes to the standard cash prize that Threadless provides for the winners. Threadless also offers challenges on creating slogans for t-shirts, but these are of minor importance and are increasingly merged with t-shirt challenges. Most challenges, apart from the continuously-running main Threadless challenge, have time limits before which submissions must be made. Typically, challenges last for a few weeks. The Challenge page for the Threadless challenge is depicted in Figure 4.13. Examples of challenges and corresponding rules are listed in Table 4.8.

Table 4.8: Some examples of challenges found on Threadless with corresponding rules and instructions (Threadless 2013)

Challenge	**Task**	**Rules**
Threadless	Your challenge: Submit a design to Threadless	You are Threadless. You make the ideas, you pick what we sell, you're why we exist. And the whole dang process starts right here. With your idea. Think you've got a show stopper up in that noggin of yours? Well pull it out of there and submit it! Check out the steps below to submit a design for all kinds of products! (Bonus points for presenting your design on more than one type of product!) The Threadless challenge is our big, ongoing, challenge. All the rest have themes, timelines, and different stuff up for grabs.
College!	Design a t-shirt inspired by one of 15 universities.	Your assignment is to create an original t-shirt design that represents one of the colleges listed above. If your design receives an A+, you could win $2500. Whether you're a current student, alum, or just a college sports fan, get creative with your favorite school's signature colors, logo, mascot, school song, or anything else that triggers your school pride. Go ahead, give it the old college try! Learn more and submit.
Onion	Create a design inspired by one of *The Onion's* headlines.	And now's your chance to win a major award in the form of cash and other swag by designing a t-shirt inspired by any headline on theonion.com. Your design can be word-free, as long as it's an interpretation of one of *The Onion's* headlines. Be sure to mention which headline inspired you in the 'about your design' section of your submission. Visit the challenge page to review the prizing and submit your design. If you think you've found a headline on The Onion and want to share it with an artist, post it below!
B&W	Submit a black & white tee design! Check out the prizing & submit	The rules are quite simple. Use black and white only in your design. You can crosshatch and halftone until you're blue* in the face but gray is not allowed! *no blue either, of course.

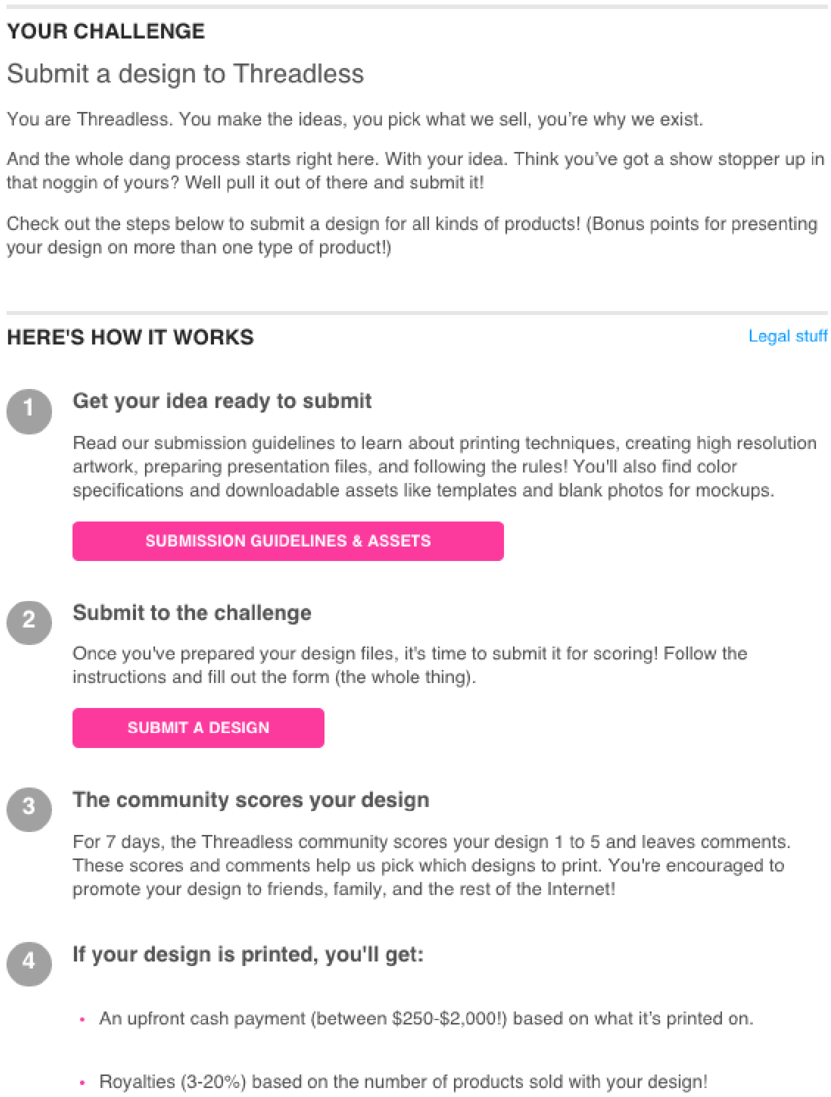


Figure 4.13: Threadless challenge page showing the instructions for participation (Threadless 2013)

In addition to design challenges, Threadless features playful competitions for fun, such as ducks in a pool lotteries, product giveaways, made-up holidays and pumpkin-carving contests. These competitions may involve social media visibility, such as posting comments to Facebook or photos to Instagram or Twitter. Winners are often selected at random and typical prizes are Threadless gift cards valued at between $25 and $50. For example, in the pumpkin-carving contest users have to turn archived designs into pumpkin carving. Painting is allowed and the resemblance to the original design can be exact or loose. To participate, a photo of the pumpkin must be posted to a Facebook contest.

Graphic design. Users take care of graphic design on their own with minimal involvement from Threadless. The site only facilitates finding collaborators and soliciting feedback from the community, neither of which is necessary for the submission of designs. People without design skills can post their ideas to challenge comments or the forum to find a willing designer to collaborate with. People use the Threadless Forum to search for collaboration partners, both in the early phases of graphic design and to finalise projects when time constraints become too pressing. It is possible to mention collaborators in the submission form in order to split the possible winnings. Unfair collaboration deals are rare and are met with passive aggression at the forum. The Threadless community organises occasionally playful competitions or challenges to encourage collaboration such as Remake Swapsies, where two designers remake each other's unprinted submissions.

Table 4.9 presents a few examples of typical graphic design processes that users have described on the website and the Threadless Forum. Graphic design begins with an idea. Old Threadless challenges and printed designs can be used as inspiration along with the usual sources artists use. The ideas for designs are selected from various sources with serendipity. The first step in realising an idea is to make one or more sketches. The first sketch is often made by hand and then scanned to graphics software for further processing. Sketches are cleaned with line work, which gets the design closer to its final shape. After that, comes colouring. During the work, the designers often seek feedback from their friends and the Threadless community. The main tools designers use are various types of pens and paper, and graphic design software. Adobe products appear to be the most popular, especially Photoshop and Illustrator. Drawing tablets make working on computers easier. Some designs feature exotic materials, such as cigarette ash, but these are fairly rare. Designers make active use of Threadless forums to solicit feedback from the community for work in progress. For many, the feedback on the forums seems to be an important part of the design process. Some users even start discussion threads where they offer to comment on work in progress. Most of the feedback is friendly, even when pointing out issues in the design. Opinions in the community appear to be split with regard to mean or unkind feedback. Some consider it unnecessarily rude and maintain that the community should be pleasant to inexperienced designers to encourage them to develop their skills, whereas others consider mean comments (especially from particular members of the community) almost as a rite of passage. They also suggest that being able to take criticism is part of the deal if one wants to do graphic design. Based on the feedback, the design is refined.

Table 4.9: Examples of graphic design process descriptions found on the Threadless website and discussion forums (Threadless 2013)

Project	Sketches	Line art	Colouring	Feedback	Refinement
The Brandywine WIP	We began this fun project by laying my vision before him with the aid of several reference images in addition to a really, really raw concept sketch by master artist. Tony soon returned with his reply and interpretation of my sketch	We continued to trade feedback daily and from the collage of process images, one can see how quickly this design moved forward from concept to a well-rounded near completion state in just seven days!	Soon after it was time to sample the often dreaded step colors	Before relaying it out there to a few fellow artists and die-hard LOTR fans for any last minute tips and/or blessings, before final touches.	The following fortnight saw us continually refine details and add minor additions/subtractions
Death by cigarettes	The design is made from two packs of cigarettes, although I only used a few cigarette butts. Didn't really smoke any of them (I quit smoking about five years ago), I just let it finish by itself and collected the ash.	Added some paper cutouts to form the eyes and nose. Dropped some cigarette ash to form the outline of the skull. The skull seemed a bit off at this point so had to tweak it a bit. Moved the nose and teeth up a bit. Changed the background to black to make it fit the shirt colour.			A bit of editing in photoshop and voila! It's done!
Bad Hare day	Since I'm going for a more loose cartoony look I went pretty straight to pencils for this design, and added some temporary type for the bottom half	I'm going to blue line the inks to print out onto Bristol board. <details> Once I've printed out my blue lines, I go to inking, using 005, 01, 05, and 08 microns, and my Pentel pocket brush pen. At this point, I scan in my Bristol board: I then clean up and separate my ink layer from the white page	Colouring. I pick a shirt color from Threadless, and colour my background layer based on that. Then I make a new middle layer, which I use to paint on the flat colors		Texturing and half toning

Submit to challenge. After a design is finalised, the submission files must be prepared according to the technical specifications provided by Threadless. The technical submission guidelines and assets are the same for different challenges. People are encouraged to read the submission guidelines (Figure 4.14) to learn about printing techniques, how to create high-resolution artwork, how to prepare presentation files and what rules to follow. Design challenges have legal terms and conditions, which must be accepted before making a submission.

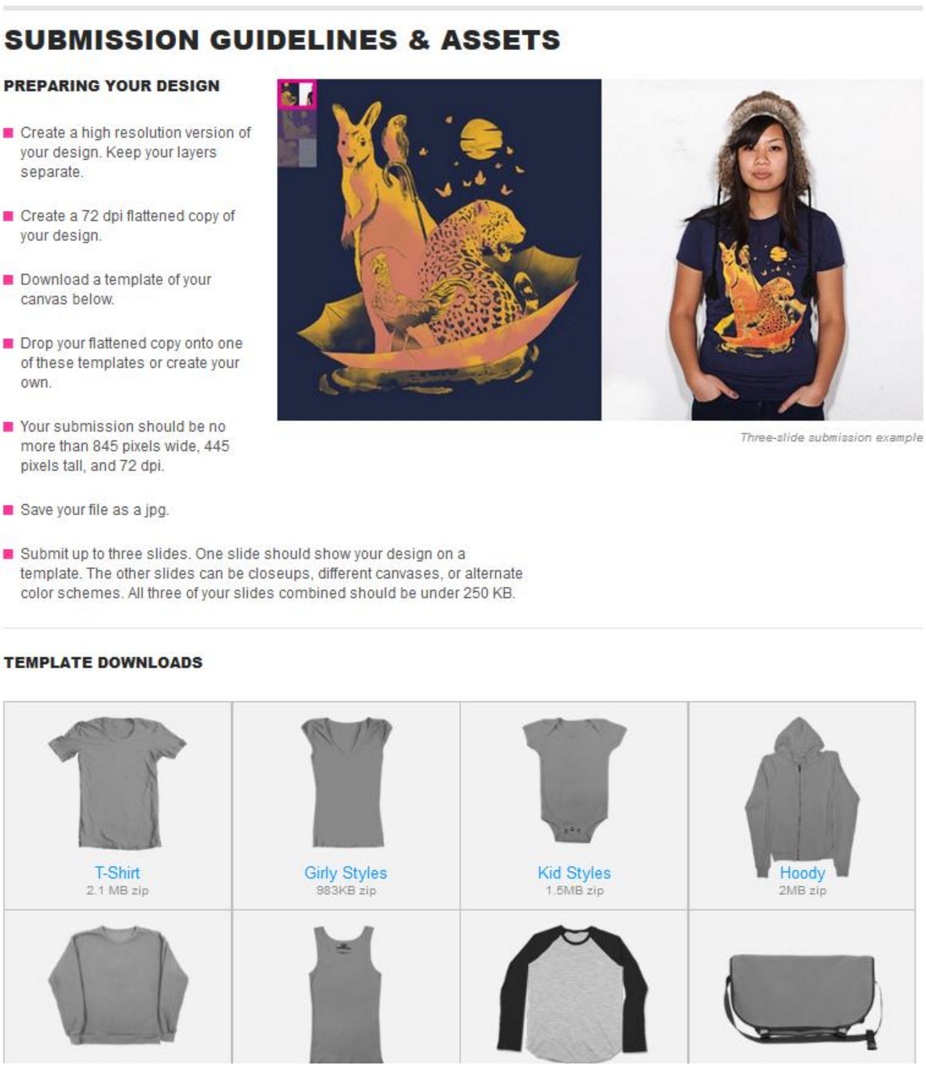


Figure 4.14: Threadless submission guidelines and assets (Threadless 2013)

Scoring. A design submitted to Threadless does not go up for scoring immediately. It is first reviewed by staff to ensure the design is not offensive. Poor quality or expectations of a low score are also reasons for Threadless to reject submissions. If a design is rejected, the designer gets an email explaining why. When a design is accepted it goes up for scoring within a couple of days and the designer gets a notification. The community usually has a week to score the design on a scale from 1 to 5. The individual scores are averaged to get the final score. There is system in place to remove the poorest designs from scoring: if after first 24 hours the score is below 1.7, the design is dropped from scoring. In a sense, Threadless uses crowds to filter out poor quality designs. The scoring page can be filtered by newest, oldest, average score, and title. To encourage scoring, Threadless organises scoring competitions and encourages users to promote their design to friends. A typical competition format is to select a $25 to $50 gift code winner randomly among people who score 50 to 100 designs in a certain challenge before a deadline. There is no reason to promote one's own work excessively, and self-promotion should be done with style. Calling all one thousand friends to vote for one's own design is not against the rules of Threadless, but it is not popular. Work in progress blog posts, on the other hand, are considered a good way to promote designs. They are also a useful for getting quick answers and feedback on details of the design. Downvoting, which refers to people giving low scores to other designs to increase their own change of getting printed, is considered particularly unacceptable. It is considered cheating by the Threadless community. Downvoting often happens in combination with voting up one's own designs, but downvoting is considered worse, because it affects the feedback from the community for the designer. The cheaters probably never win because the ultimate printing decision is made by Threadless staff. The key to winning is making a good design, not getting the highest score. On the other hand, Threadless states that 'votes are so important to our decision in picking designs for printing we need more people voting' (Threadless 2013). Voting on designs is a major activity on the website, in addition to surfing the forum. It is an easy and somewhat pleasant activity that can even induce mild flow experience. Most designs are of decent quality and some are excellent. After voting on a design, a user is moved automatically to the next design, creating a temptation to vote on just one more design in the expectation that perhaps the next one will be a rare gem. After a hundred designs, voting can get tedious. According to comments on the discussion forum, waiting for voting results can be a nerve-wrecking experience for designers. Vote scores provide feedback to designers. Good scores may raise hopes of getting printed, whereas low scores can feel crushing, especially if cheating or downvoting is suspected.

Winners. Designers whose designs get printed at Threadless typically receive $2,000 in a cash and $500 gift card. Special challenges offer extra prizes, such as tickets to events, gift cards to other stores, baseball bats, or higher amount of cash, such as $3,000 or $5,000. The understanding in the community seems to be that the focus of designers should not be on winning, which is rare, but on making stuff and enjoying it.

4.3.4 User experience at Threadless

The Threadless process appears to be quite simple and straightforward from the user's viewpoint. User activities are clearly displayed on the front page: make, pick, play or shop. Make refers to graphic design, which takes place mostly outside the website, at participants' homes and offices. Pick means scoring and commenting on the designs currently in evaluation. Play involves all the other activities possible on the site, most importantly the Threadless Forum, the meeting place for the Threadless community. Prominent topics on the forum are art and design, and the promotion of design projects in progress. Other activities in this category include the staff blog that features short news items related to the company, and Tee V, a site dedicated to short videos. The company also has comprehensive coverage on various social media sites. Shop is self-explanatory. Being nice to other people is emphasised. Even the headline for commenting on designs reads 'Say somethin' nice' (Threadless 2013). The site feels active, and the rate at which new designs appear is impressive. Learning new skills and improving as a graphic designer appear to be important motivational factors for creating and submitting designs to Threadless. As one user recounts on the Threadless Forum, many people, are drawn to the site by the promise of money. Disappointment may ensue if these promises are not realised. After all, most people's designs perform poorly in scoring and a vast majority of the designs will never get printed. A different perspective may make the experience much more pleasant for the participants. When submitting designs is not about getting printed, but about improving as an artist, it is possible to keep submitting designs, have fun, and develop one's artistic skills despite the lack of financial success. This particular user further explains that he submitted about 30 designs to the site before getting one printed. According to him, this is typical in his network of designers. Several other users mention learning and career advancement as motivational factors: existing research corroborates the observation (Brabham 2010). In addition to creating and submitting designs, the site offers other possibilities for learning, such as tutorials on graphic design tools and process blogs that describe the creation process of particular designs in detail. Both these learning resources are mostly user-generated. A major part of the value Threadless provides for the user is the community of other users and designers. The community helps in the design process and provides feedback and support to its members. Users also organise playful challenges every now and then on the forum, such as Swapsies and ThreadDuels. These little ideas help keep Threadless fresh.

Due to my lack of graphic design skills, I did not submit designs to Threadless. Instead I focused on scoring designs and surfing the discussion forum. Scoring designs is easy and fun, and somehow the most obvious thing to do. Scoring is also the easiest way to get something done if you cannot decide what else to do. Usually, I scored designs using the default settings, which meant that I viewed the designs in submission order. Most designs are of a decent quality and some are particularly striking. Most of the time, I preferred single design mode when scoring designs, which made it easy to continue scoring designs. In single design mode, the platform presents a new design automatically after scoring, which tempts one to continue scoring just one more design, and so on. I found scoring to be a pleasant activity, which induced mild flow in single design mode. The flow experience seems to be related to the appearance of new tasks at a suitable rate and the expectation of striking designs. I still wish the system was able to load a few designs on the computer in advance. There is always a lag between scoring a design and getting a new one. I only used the multi-design scoring mode once when the Internet connection was slow. The Threadless challenge was my favourite for scoring. I entered the occasional contest to win a t-shirt if it was particularly attractive. After around hundred or so designs, scoring becomes tedious, especially if all the designs were from the same challenge. Randomisation might help to alleviate this issue by adding more variety. I participated in scoring challenges when I saw them, but they did not change my behaviour. During the observation period I would have scored designs in any case, but the contest directed me to challenges that I might otherwise have ignored. I think that the accuracy of my evaluations varied considerably. As I am a critical evaluator, I believe there is nothing wrong as long as the opinion is honest.

Scores provide feedback to designers; according to forum discussions, waiting for the results of scoring can be a nerve-wrecking experience. Good scores may raise the hopes of getting printed, whereas low scores can feel crushing, especially if cheating or downvoting is involved in the challenge. There is evidence of occasional cheating and downvoting on Threadless. In these cases, a user creates multiple accounts to give his designs high scores and everyone else low scores, effectively determining the results of scoring. This is very annoying for other participants, particularly because it ruins community feedback. There is a common belief that cheaters will not win, because the final decision on what to print is made by Threadless.

5 Cross-case analysis and results

In this chapter. cross-case analyses are performed. As the concept of collective intelligence is still somewhat fuzzy, it is necessary to take into account several different interpretations of the phenomenon revealed in the literature review. Observed patterns gleaned from case descriptions are matched to theoretical patterns derived from different theoretical frameworks, progressing from general descriptions to more detailed examinations. This theory and method triangulation helps to increase the study's construct validity.

5.1 Collective intelligence genome

As a first analytical step, the phases where the crowd contributes to the innovation process are identified. This theoretical lens describes the division of work and its rough nature throughout the innovation processes encountered in the study. The case sites are described using a framework for classifying the building blocks of collective intelligence systems called the genome of collective intelligence, which was developed by Malone et al. (2010).Table 5.1, Table 5.2 and Table 5.3 present the collective intelligence genomes for the three cases. The answers to key questions are interpreted from the case descriptions provided in the previous chapter. Although some of the cases have been discussed in the original paper (Malone et al. 2010), the use of ethnographic methods in this study allows a more detailed analysis.

Table 5.1: Collective intelligence genome for OpenIDEO

OpenIDEO Phase	What		Who	Why	How
Challenge	Create	Challenge brief	Hierarchy	Extrinsic	Hierarchy
Inspiration	Create	Inspirations	Crowd	Intrinsic	Collection
Synthesis meeting	Decide	Themes	Hierarchy	Extrinsic	Hierarchy
Concepting	Create	Concepts	Crowd	Intrinsic	Collection
Applause	Decide	Most applauded, viewed and commented concepts	Crowd	Intrinsic	Voting
Shortlist selection	Decide	Shortlist	Hierarchy	Extrinsic	Hierarchy
Refinement	Create	Refined concepts, prototypes, visualisations	Crowd	Intrinsic	Collaboration
Evaluation	Decide	Evaluation of concepts	Crowd	Intrinsic	Voting
Winning concepts	Decide	Challenge winners	Hierarchy	Extrinsic	Hierarchy
Realisation	Create	Prototypes, tests, business models, implemented solutions	Crowd	Intrinsic	Collaboration

The collective intelligence genome for OpenIDEO reveals an alternating pattern between activities carried out by the hierarchy and the crowd. Phases are focused on creation, on the one hand and decisions, on the other. The hierarchy provides instructions, based on which the crowd then creates contributions. The hierarchy makes all the final decisions, perhaps influenced by crowd decisions. All crowd decisions are aggregated. The process is largely about creating collections and then filtering out the best contributions. As OpenIDEO does not offer any rewards to the participants, the motivation of the crowd is mostly intrinsic by nature.

Table 5.2: Collective intelligence genome for Quirky

Quirky Phase	What		Who	Why	How
(Challenge)	Create	Design brief	Hierarchy	Extrinsic	Hierarchy
Idea submission	Create	Product concept	Crowd	Extrinsic	Collection
Community curation	Decide	Popularity of concepts	Crowd	Extrinsic	Voting
Product evaluation	Decide	Winning concepts	Hierarchy (+ crowd)	Extrinsic	Hierarchy / Voting
Research	Decide	Answers to research questions	Crowd (+ hierarchy)	Extrinsic	(Consensus?)
Design, submission	Create	Design suggestions	Crowd (+ hierarchy)	Extrinsic	Collection
Design, voting	Decide	Popularity of design suggestions	Crowd (+ hierarchy)	Extrinsic	Voting
Branding, submission	Create	Names and taglines	Crowd (+ hierarchy)	Extrinsic	Collection
Branding, voting	Decide	Popularity of names and taglines	Crowd (+ hierarchy)	Extrinsic	Voting (ELO compare two)
Engineering + Finalisation	Create	Final product design	Hierarchy	Extrinsic	Hierarchy
Market research (pricing)	Decide	Price	Hierarchy (+ crowd)	Extrinsic	Voting
Green light	Decide	Commercialisation	Hierarchy	Extrinsic	Hierarchy
Manufacturing	Create	Products	Hierarchy	Extrinsic	Hierarchy
Sales			Hierarchy	Extrinsic	Hierarchy

The first thing to note about the collective intelligence genome for Quirky is the heavy reliance on hierarchy. Apart from the initiation of the process, the crowd plays only a relatively small role compared, for example, with OpenIDEO. Although the crowd creates the ideas and makes an initial assessment, after that, the hierarchy plays a major role in the development process, with frequent but minor inputs from the crowd in predefined phases. Lots of effort is placed on filtering at the beginning. Later, the focus is on refinement through decisions on details of the project. Quirky relies on extrinsic motivation by using its proprietary algorithm to measure and reward the contributions of each participant to the development projects.

Table 5.3: Collective intelligence genome for Threadless

Threadless Phase	**What**		**Who**	**Why**	**How**
Challenge	Create	Design challenge	Hierarchy	Extrinsic	Hierarchy
Graphic design (Idea, collaboration, sketch, line art, colouring)	Create	T-shirt design	Crowd	Intrinsic	(Collaboration)
Feedback	Decide	How to improve the design	Crowd	Intrinsic	Consensus
Refinement	Create	T-shirt design	Crowd	Intrinsic	(Collaboration)
Submit	Create	Collection of designs	Crowd	Intrinsic/ Extrinsic	Collection
Review	Decide	Acceptance to scoring	Hierarchy	Extrinsic	Hierarchy
Scoring	Decide	Average score of designs	Crowd	Intrinsic	Averaging
Winners	Decide	Winning designs	Hierarchy	Extrinsic	Hierarchy
Print	Create	Final product	Hierarchy	Extrinsic	Hierarchy

The most interesting thing about the genome of collective intelligence for Threadless is the large role the crowd plays in the early phases of the process. Apart from the creation of challenges by the hierarchy, the crowd works almost on its own until the very end of the process. The crowd takes care of the creation and initial evaluation of the designs: the hierarchy then picks the best for production. Despite the monetary rewards for selected designs, the motivations of the crowd appear to be mostly intrinsic: the possibilities for financial gain are only two of the five main motivations identified in previous research (Brabham 2010).

Comparison of the three collective intelligence genomes reveals differences in emphasis on phases of the process and division of labour. Three distinct patterns emerge: a crowd-heavy approach on Threadless, a hierarchy-heavy approach on Quirky and an alternating one at OpenIDEO. Another difference is in the emphasis of the activities. OpenIDEO is highly focused on the first parts of the generic innovation process, namely understanding the problem and generating, refining and evaluating solution suggestions. Design concepts are the end result of the process. Quirky, on the other hand, begins where OpenIDEO left off. The first inputs to the system from the participants are design concepts, which are then evaluated and developed to products mainly by Quirky staff members. Threadless staff members define the problem (design challenge) at the beginning and, after that, only become involved again at the end of the process. If the interaction on the Threadless Forum is not taken into account, the site is involved in the development process only after the designs are finalised. Figure 5.1 visualises the differences between the crowdsourcing sites. Despite the surface similarities between the cases, they are all implemented quite differently.

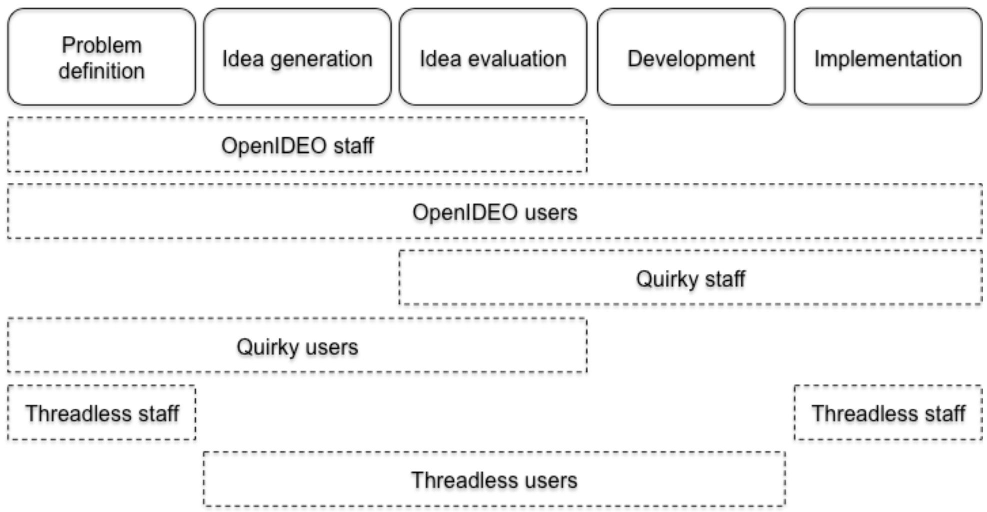


Figure 5.1: Distribution of work on investigated crowdsourcing sites.

5.2 Crowdsourcing systems as inventors

If crowdsourcing leads to collective intelligence similar to the swarm intelligence of social insects, it may be that participants of a crowdsourcing effort collectively function similarly to an individual inventor. In the second theoretical lens, the creation of innovations is framed as a recombinant search over the space of possible solutions (Fleming and Sorensen 2001). The space of possibilities is naturally enormous; accordingly, it is not possible to try out every solution. Instead, some other approach must be used. While search algorithms on computers often rely on approaches such as local hill-climbing (Russell and Norvig 1995), this is not how humans typically create inventions. Rather, the individual inventors rely on more flexible search heuristics (Maggitti et al. 2013). Under the paradigm of innovation as search, it is assumed that searching for solutions locally (i.e., focusing on the neighbourhood of solutions already familiar to the searching agent) is easy compared to a distant search (searching for solutions not related to the current knowledge of the searching agent). Afuah and Tucci (2012) propose that problem solving through crowdsourcing works because it transforms the distant search to a local one. Crowdsourcing can therefore be considered as some kind of intelligent search algorithm. But how does that algorithm work, and how does it compare to the heuristics used by an individual inventor? The following analysis extends the pilot study reported in Salminen (2013b).

Crowdsourcing can improve problem-solving performance in cases where distant knowledge is needed (Afuah and Tucci 2012). Instead of performing a distant search, the problem solver may crowdsource the task to a large number of participants, all of whom perform only local searches. If the number and diversity of participants are high enough, the chances are that someone is already in the neighbourhood of the solution, and can find the solution by a local search. Thus, distant search is transformed into local search and the performance of problem solving is improved. The question arises as to whether a search process using crowdsourcing resembles the search process of individual inventors. Should we consider a crowdsourcing system as a single searching entity or a collection of individual inventors? To explore this question, the innovation processes of the case companies are compared to the inventor's search process stages (Maggitti et al. 2013).

Table 5.4: OpenIDEO as an inventor.

Inventor's search stage	**OpenIDEO process**	**Notes**
Stimulus	Challenge design	Preparation of challenge brief
	Challenge brief	
Net casting	Inspiration	Crowd submits solutions that have worked elsewhere in similar situations
Categorising	Synthesis meeting	OpenIDEO staff members organise the inspirations
Linking	Concepting	Crowd creates and submits concepts based on the challenge brief and organised inspirations
Discovery	Applause	Crowd votes and comments on submitted concepts
	Shortlist selection	OpenIDEO staff members select 20 best concepts for refinement
	Refinement	Crowd members improve on shortlisted concepts
	Evaluation	Crowd evaluates shortlisted concepts on multiple criteria
	Winning concepts	OpenIDEO staff members select winning concepts
	Realisation	Anyone is allowed to implement the winning concepts

In Table 5.4 the search process of individual inventors is compared to the OpenIDEO challenge phases and user experience during the Concepting phase. Although the phases and stages are not divided up in the same way, the content appears to fit. OpenIDEO challenges proceed roughly according to the search model of an individual inventor. Therefore, OpenIDEO could be considered to be functioning as a single inventing entity. However, the picture changes slightly when the activities of users in the Concepting phase of the process are investigated in more detail. Table 5.5 compares user experience during the Concepting phase to the inventor's search process.

Table 5.5: User experience during the OpenIDEO Concepting phase compared to the inventor's search process.

Inventor's search stage	User experience during Concepting phase
Stimulus	Reading challenge brief and instructions, watching videos
Net casting	Reading inspirations and concepts, collecting pieces for a concept
Categorising	Trying to figure out the big picture, sleeping
Linking	Coming up with an idea, creating a concept
Discovery	Prototyping, developing a concept, submitting a concept, updating a concept

Interestingly, the user experience during the Concepting phase of OpenIDEO challenges also matches the individual inventor's search process. Thus, the search process of OpenIDEO can be described as a search-within-search. The process resembles the search process of an individual inventor at the system level. However, during the Concepting phase, each individual participant needs to go through the individual inventor's search process. It does appear, however, that OpenIDEO has taken steps towards a system that could be considered a single searching agent or inventor, although it is not there yet.

Table 5.6: Quirky as an inventor.

Inventor's search stage	Quirky process	Description
Stimulus	(Challenge)	Members of Quirky staff may create a challenge brief
Net casting Categorising Linking	Idea submission	Crowd members create and submit product ideas
Discovery	Community curation	Crowd members vote and comment on submitted ideas
	Product evaluation: UC	Members of Quirky staff put ideas under consideration
	Product evaluation: Preval	Members of Quirky staff select ideas for live evaluation
	Product evaluation: Live eval	Members of Quirky staff select winning ideas
	Research	The crowd answers a survey created by Quirky staff
	Design	The crowd creates, submits and votes on design suggestions
	Branding	The crowd submits and evaluates names and taglines
	Engineering & finalisation	Members of Quirky staff finalise the product design
	Market research: Launch, Pricing, Pre-order	The crowd estimates prices for products
	Green light	Members of Quirky staff decide on manufacturing
	Manufacturing	Outsources manufacturing
	Sales	Sales through website and other channels
	Cash in	The crowd gets paid according to influence scores

Table 5.6 presents a similar comparison between the inventor's search stages and the Quirky process. The most striking observation is that most of the inventor's search process stages are taken care of solely by the crowd. The Quirky platform and staff become involved in the development process only at the discovery stage. This observation is in line with how Quirky itself describes its approach: the stated goal of the company is to provide a route to market for inventors, who would not otherwise commercialise their inventions. It therefore makes sense that Quirky is mostly concerned about what happens after an individual inventor has made an invention. The Quirky platform does not function as a unified inventor, but instead chooses to carry on the development work of a few, selected inventors. Inventing is still done by single individuals.

Table 5.7: Threadless as an inventor.

Inventor's search stage	**Threadless process**	**Description**
Stimulus	Challenge	Members of Threadless staff create a challenge brief
Net casting	Idea	Crowd members come up with ideas for designs
Categorising		
Linking	Graphic design (sketches)	Crowd members create one or more drafts of the design idea
Discovery	Feedback	Feedback from Threadless community and/or other sources
	Graphic design (refinement)	Crowd members finalise graphic designs
	Submit	Crowd members submit designs
	Review	Members of Threadless staff review submitted designs
	Scoring	Crowd members score designs
	Selecting winners	Members of Threadless staff select winning designs
	Printing	Threadless prints winning designs

Comparison between the inventor's search stages and the Threadless process is presented in Table 5.7. The pattern is similar to the one revealed in Quirky: the crowd performs most of the inventor's search stages. The crowdsourcing platform facilitates the discovery phase by providing mechanisms for gathering feedback (the discussion forum) and evaluating designs. The manufacturing of the finished design is the responsibility of the company.

If creation of innovation is a search problem, then what kind of search algorithm is crowdsourcing? The analysis reveals a common pattern: on all sites, the individual participants need to go through most of the inventor's search process on their own. On Quirky and Threadless, this is easy to see as the companies become seriously involved with the development process only at the discovery phase. At OpenIDEO, the picture is more complex. The stages of an individual inventor's search process can be identified both at the system level and in the activities of the participants. The OpenIDEO process roughly follows the phases of the inventor's search process. In the Concepting phase, however, the individual participant has to go through all phases of the process alone. Therefore, the search algorithm used on OpenIDEO can be described as a search-within-search, whereas Quirky and Threadless explicitly collect and evaluate solutions found by individual inventors. All the sites appear to be most appropriately described as consisting of collections of individual inventors, with platforms providing varying levels of support to search activities. A question arises as to whether it is possible to design a crowdsourcing system that could function as a single inventing entity, without the need for the individual participants to perform all the stages of the inventor's search. Differences in focus areas are found here too, corroborating the findings from the first theoretical lens, the genome of collective intelligence.

5.3 Collective intelligence systems

The third theoretical lens is specifically concerned with the emergence of 'more than the sum of the parts' at the global level of a system solely as a result of interactions at the local level. Schut (2010) gives a framework for evaluating these kinds of collective intelligence systems. There are three enabling and five defining properties of these systems. If enabling properties are observed, the system can be a collective intelligence system. If defining properties are also in place, the system can be called a collective intelligence system. Assessment of these properties on case sites are presented in Table 5.8, Table 5.9 and Table 5.10.

Table 5.8: The enabling and defining properties of collective intelligence at OpenIDEO

OpenIDEO	**Enabling properties**			**Defining properties**					
Phase	**Adaptivity**	**Interaction**	**Rules**	**Local**	**Global**	**Emergence**	**Randomness**	**Redundancy**	**Robustness**
Challenge	Different challenges	Unknown	Design brief		No	No	No	NA	NA
Inspiration	Different challenges, challenge-specific missions	Reading content, comments, submissions, applause	'Do NOT post your ideas yet!'	Inspirations, views, comments, applause	Collection of inspirations	Collection, sum of the parts	Random sorting of submitted inspirations	Hundreds of users	Results do not depend on single user
Synthesis meeting	Unknown	A meeting	Unknown	Inspirations	Themes	From inspirations to themes	Unknown	Meeting	Single meeting, but several participants
Concepting	Different challenges, challenge-specific themes	Reading content, comments, submissions, applause	'Concepts are here to build upon' 'Be visual'	Concepts, views, comments, applause	Collection of concepts	Collection, sum of the parts	Random sorting of submitted concepts	Hundreds of users	Results do not depend on single user
Applause	Different challenges	Reading content, comments, applause	'Inspirations, comments and concepts can be applauded'	Views, comments, applause	Collection of sorted concepts	From solo applause to sorted list of concepts	Random sorting of submitted concepts	Hundreds of users	Results do not depend on single user
Shortlist selection	Unknown	A meeting	Unknown	Concepts	Shortlisted concepts	No	Unknown	Single meeting	Several participants
Refinement	Different instructions on improvement	Reading content, comments, applause	'Refining is as much about adding more detail as it's adding clarity'	Views, comments, applause, concept updates	Collection of refined concepts	No, unless discussion counts	No	Hundreds of users	Results do not depend on single user
Evaluation	Evaluation questions depend on the challenge	Reading content, comments, applause, evaluations	'Evaluate all shortlisted concepts'	Individual evaluations	Aggregated evaluations	Sum (or average) of parts	No	Hundreds of users	Results do not depend on single user
Winning concepts	Unknown	A meeting	Unknown	Shortlisted concepts	Winning concepts	No	Unknown	Meeting	Several participants
Realisation	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable

5.3.1 OpenIDEO as a collective intelligence system

Table 5.8 summarises the characteristics of OpenIDEO from the perspective of collective intelligence systems. All the enabling properties of collective intelligence systems can easily be observed at OpenIDEO. The enabling properties are mostly concerned with the agents in the system. In this and other investigated cases, the participants are humans, and it is safe to assume that people are adaptive. The implication is that the system is also adaptive. This point is important, because Schut's (2010) framework is concerned with all kinds of agents, and it is not always necessarily the case that agents are adaptive. The OpenIDEO website offers possibilities for many types of interaction, including viewing the content, submitting content (inspirations and concepts), commenting and applauding. Behind the scenes, live meetings take place. The interaction resembles common web forums. The meetings presumably consist of normal human interaction. Activities on the website are governed by the cultural norms of the community. In addition, the site features various explicit rules, such as challenge briefs and instructions for specific innovation process phases.

At OpenIDEO, the local interactions of the agents are typical activities that users can usually carry out on web forums: viewing content, commenting and applauding. At the global level, the system produces collections of submitted content that can be ordered by the number of views and comments or the amount of applause. Individual submissions and evaluations are simply aggregated together. Although difficult to identify, emergence does not appear to play a major role in the functioning of the site. Emergence is limited to consensus on inspiration themes emerging from the synthesis meeting through normal face-to-face discussion, with content sorted by the aggregated evaluations. The collection of concepts is simply a sum of its parts, the individual concepts. Being able to sort them by views, comments or applause adds only a little compared to the cognitive work needed to produce the individual concepts. A certain amount of randomness is naturally inherent in websites. The behaviour of users is not deterministic; inspirations and concepts are sorted in random order by default. The number of participants is relatively large. As a result, the functioning of the site is not dependent on individual users. Although a particular contribution might be lost, the site would still work much the same even if some users were lost.

Table 5.9: The enabling and defining properties of collective intelligence at Quirky

Quirky	**Enabling properties**			**Defining properties**					
Phase	**Adaptivity**	**Interaction**	**Rules**	**Local**	**Global**	**Emergence**	**Randomness**	**Redundancy**	**Robustness**
(Challenge)	Different challenges	Unknown	Unknown		No	No	No	No	No
Idea submission	Large range of different ideas	Submission	'Idea description is limited to 140 characters'	Idea submissions	Collection of submitted ideas	Collection, sum of the parts	No	Thousands of users	Results do not depend on single user
Community curation	Different products	Votes, comments, updates	15 votes per day	Votes, comments	Most active ideas	Sorting of ideas by votes (not available for users)	Large amount of new content	Thousands of users	Results do not depend on single user
Product evaluation	Different categories and products	Discussion, chat, voting, sentiment meter, etc.	'Selected ideas get designed and likely commercialised'	Individual opinions	Consensus	Consensus on which ideas get designed and commercialised	No	Dozens of participants	Results do not depend on single participant
Research	Different questions, different products	Usually survey	'Answer all research questions'	Individual answers	Aggregated results	Aggregated results	No	Thousands of users	Results do not depend on single user
Design, submission	Different tasks, different products	Design submissions	'Each community member is allowed to submit one concept'	Design submissions	Collection of designs	Collection, sum of the parts	No	Thousands of users	Results do not depend on single user
Design, voting	Different products	Voting, commenting	'…you've given certain number of votes…'	Voting, commenting	Aggregated results	Sorting of concepts by votes (not available for users)	Random order	Thousands of users	Results do not depend on single user
Branding, submission	Different instructions, different products	Name or tagline submissions	'Don't include the word Mac'	Name of tagline submissions	Collection of suggestions	Collection, sum of the parts	No	Thousands of users	Results do not depend on single user
Branding, voting	Different products	Voting (new system)	'You'll earn valuable influence after voting on 20 names'	Voting	Aggregated results	Sorted names or taglines	Random order of names during voting	Thousands of users	Results do not depend on single user
Engineering + Finalisation	Different products	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
Market research (prising)	Different products	Pricing game	'Share your pricing insights within 7 days…'	Price estimates	Aggregated prices	No, aggregated (averaged) prices	No	Thousands of users	Results do not depend on single user
Green light		Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
Manufacturing		Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
Sales		Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown

5.3.2 Quirky as a collective intelligence system

A summary of the enabling and defining properties of collective intelligence systems for Quirky is presented in Table 5.9. As with OpenIDEO, the participants at Quirky are human beings. They are a natural source of adaptivity, which is enough to make the system as a whole adaptive. The system adapts to different phases of the development process: there are different tasks depending on the phase and product under development. The system is able to develop different kinds of products, from kitchen utilities to electronics. The interaction possibilities are similar to those found at OpenIDEO: submissions, reading and writing comments, and participating in various voting schemes. Direct messages between participants and forum discussions are also possible. The specialty of the site is streamed and recorded videos from decision-making, brainstorming and feedback sessions in which viewers can participate through the chat interface. Rules are defined somewhat more explicitly than at OpenIDEO. This is especially true for the measurement of contributions to product development in order to calculate the influence scores for each participant. Specific tasks have their own instructions and limits, such as character count limits for submissions, or the number of votes for idea evaluation. Cultural norms are naturally also present, implicitly guiding the behaviour of participants. Occasionally these norms are made explicit when users attempt to correct the behaviour of other participants who appear to have broken some of the norms.

The distinction between local and global levels is most easily observed in the creation of collections of submissions and the aggregation of individual evaluations and survey answers. Quirky does not show these aggregated results directly to participants while the evaluation is still underway. Instead, the ideas can be organised by their activity. It is not clear exactly how an activity is measured. It is supposedly about the rate of views, comments and votes on ideas. At the local level there are individual actions, such as views, comments and votes, and at the global level there are ideas sorted by their activity levels. Two types of emergence can be identified at Quirky: the aggregated results of individual evaluations or other contributions, and consensus formation during the idea evaluation process conducted by the Quirky staff (Under Consideration, Preval and Live Eval). These present themselves as consensus about which products to develop and commercialise, sorted content (for example the most active ideas), and aggregated results of surveys, naming games and pricing games. Randomness comes to the site through the large number of participants and the rate of new content creation this entails. Participants are exposed to new content semi-randomly due to this high rate of new content. As the participants are not deterministic, there is an element of randomness that can create critical behaviour. For instance, it is random which ideas end at the top of the most active ideas list. However, once ideas have reached that list, they may gather many more votes in a short period of time due to feedback loop: the most active ideas get more attention, which makes them even more active. One of the participants described this phenomenon at the Quirky discussion forum (Quirky 2013) as ‘once you get on the Most Active Roller Coaster, you keep riding!’ Redundancy and robustness are again based on the large number of participants (tens of thousands, according to Wang (2011)). In relation to the mass of participants, it is not very important what a few do, as long as the participants who provide high quality solution suggestions do not disappear completely.

5.3.3 Threadless as a collective intelligence system

Threadless is described in the framework of collective intelligence systems in Table 5.10. With regard to the enabling properties of collective intelligence systems, the picture at Threadless is very similar to the two previous cases. The participants are people and they provide a major element of adaptivity to the system. Again, the interaction resembles typical interactions on a web forum. However, it should be noted that on Threadless an actual web forum is the stage for a significant part of the activities. This forum is the place for the community to gather to exchange opinions, share ideas and give feedback on designs in progress. Other important activities are design submissions and scoring and commenting on designs. Threadless features two sets of rules. The cultural norms are made partially explicit, especially by the promotion of the instruction to be pleasant, which appears to be the main rule for Threadless. Another set of rules consists of detailed instructions for specific tasks, such as challenge briefs, submission guidelines and scoring rules. In addition, the activities enabled by the website design can be considered as a third implicit set of rules.

The clearest place where a distinction between local and global levels can be made at Threadless is the distinctions between individual scores given by the participants and the aggregated average score. Another place where differences between local and global levels may exist is in the creation of the designs themselves. However, this activity does not really take place on the site but at the homes and offices of the individual participants. Emergence might take place at the discussion forum as a consensus on how to refine a design emerging from discussions. Final designs could also be considered as an emergent result of the design process but this again is mostly conducted outside the site, in the work of a single designer, and is therefore outside the scope of this study. Similarly to the other cases, lists of sorted designs emerge during scoring. While individual participants score only some of the designs, aggregation of the scores results in globally sorted designs at the system level. Again, randomness features through inherent randomness in the behaviour of the website users. Other sources include random exposure to new forum messages and designs. Redundancy and robustness are the same as in OpenIDEO and Quirky: the large numbers of participants create the redundancy that makes the system robust against single defective users. A particular t-shirt design may be missed, but there are always plenty of other options to choose from; as long as the site attracts large numbers of participants, it will not run out of design suggestions.

Table 5.10: The enabling and defining properties of collective intelligence systems at Threadless

Threadless	**Enabling properties**			**Defining properties**					
Phase	**Adaptivity**	**Interaction**	**Rules**	**Local**	**Global**	**Emergence**	**Randomness**	**Redundancy**	**Robustness**
Challenge	Different challenges	No	Unknown		No	No	No	No	No
Graphic design (Idea, collaboration, sketch, line art, colouring)	Different designs and challenges	Forum discussion	Be nice, challenge brief	Messages, sleeping, thinking, sketching, etc.	Design	Design?	No	Thousands of users	Results do not depend on single user
Feedback	Different feedback	Forum discussion	Be nice	Reading and commenting on the forum	Discussion	Consensus on refinement	Semi-random exposure to newest posts on the forum	Thousands of users	Results do not depend on single user
Refinement	Different designs	No or possibly private	No	Changes to design	Improved design	Final design?	No	Thousands of users	Results do not depend on single user
Submit	NA	Submission form	Submission guidelines	Design submission	Collection of designs	Collection, sum of the parts	No	Thousands of users	Results do not depend on single user
Review	Different challenges	Unknown	Unknown	Designs	Accepted designs	No	Unknown	No	No
Scoring	Different designs, different challenges	Scoring, comments	'For 7 days Threadless community scores your design and leaves comments'	Individual scores, comments	Aggregated scores	Sorted designs	Semi-random exposure to newest designs	Thousands of users	Results do not depend on single user
Winners	Not applicable	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
Print	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable

5.3.4 Crowdsourcing sites as collective intelligence systems

The investigated cases satisfy all the enabling criteria of collective intelligence systems. People, the agents of the system, can adapt to different situations. The adaptivity of participants is sufficient for the system as a whole to be adaptive. It is probably the most important source of adaptivity in the investigated systems. As a result, the crowdsourcing systems are able to produce different outputs at different phases of the innovation processes. Interactions are typical for web forums. Participants can submit various contributions and comments, read what others have posted and participate in tasks, such as idea evaluation and surveys. Although not entirely explicit, there are rules, including cultural norms that communities are supposed to follow in each of the crowdsourcing platforms. Overall, the enabling criteria are easy to observe in the investigated cases. The enabling properties are mostly concerned with the local or agent level. Here the agents are humans. Humans are adaptive, they interact with each other, and they follow implicit and explicit rules, such as cultural norms, challenge briefs and task instructions. Most systems involving interacting humans would probably fulfil the enabling properties of collective intelligence systems.

Of the properties that define collective intelligence systems, the satisfaction of randomness, redundancy and robustness criteria are also easily observed. In addition to the randomness inherent in non-deterministic human behaviour, the sites often display the content in randomised order. The cultural rules and the tasks are shared between the many individuals participating in activities on the site. Hundreds or thousands of participants create redundancy, which makes the systems robust against failures such as some of the participants leaving the site. One particular piece of content may be lost, but another piece of content will replace it. Results do not depend on a single user. Randomness, redundancy and robustness can probably be found on almost any interactive website. These properties are therefore not very interesting from the viewpoint of identifying collective intelligence. The distinction between local and global levels and emergence are the defining factors of collective intelligence at crowdsourcing sites. These two criteria turned out to be somewhat challenging to identify. In the investigated cases, local or agent level behaviour consists of creating and submitting content, and viewing, commenting and voting on the content that other participants have submitted. At the global or system level, there are collections of submitted content, aggregated evaluations and consensus decisions. Potential emergents on the sites are consensus emerging from live or web forum discussions, collections of content emerging from individual submissions, and crowd opinion emerging from aggregated evaluations. Table 5.11 summarises the identified emergents. It appears that there is not a huge difference between the intelligence of the systems compared to the intelligences of the participants within them. This observation is in contrast to the swarm intelligence of social insects, where the individuals are dumb but the system is smart in comparison. However, emergence is a convoluted concept. Due to the difficulty of identifying emergence in practice, there could still be other, unidentified types of emergence present at the case sites.

Table 5.11: Summary of identified sources of emergence at the investigated crowdsourcing sites

Local	**Global**	**Emergence**	**Notes**
Comments during a meeting or forum discussion	Discussion	Consensus	Group behaviour in general is outside the scope of this research
Content creation and submission	Collection of content	Collection	Merely the sum of the parts
Individual evaluations	Aggregated evaluations	Crowd opinion	

Potentially intelligent group behaviour during discussion is outside the scope of this study; further discussion of the topic is left to psychologists and social psychologists. The amount of emergence in the creation of collections is questionable, as they are by definition sums of their parts. Aggregated evaluations that make it possible to sort those collections based on the crowd's collective opinion show more promise as an emergent, collectively intelligent phenomenon. Aggregated evaluations can be used to create order that was not there before. Even if each participant sees only a small amount of the total content, order among all the content emerges through the aggregation of evaluations. A simple example is the sorting of designs by average scores, as is done on Threadless, or by activity, as is done at Quirky, or by views, comments or applause, as is possible at OpenIDEO. If there is something that could be described as collective intelligence at the sites, the aggregated evaluations are perhaps the best candidates. The question is, how intelligent is the crowd when producing these evaluations? How does it compare to individual decision makers?

5.4 Wisdom of crowds

The fourth theoretical lens is focused on the output of the system as a statistical or probabilistic phenomenon. The following analysis extends the work started in Salminen (2014). One of the premises of crowdsourcing is that it exploits the wisdom of crowds. Under the right circumstances, the aggregated judgment of a crowd can be closer to the truth than that of the best individuals in the crowd. To be wise, a crowd should be diverse (Page 2007), judgments of its members should be independent, and there should be a way to aggregate the judgments (Surowiecki 2005). Previous research has shown that diverse problem solvers can outperform the best problem solvers both in simulations (Hong and Page 2004) and experimental settings (Krause et al. 2011). On the other hand, even a minor social influence can decrease the accuracy of a crowd (Lorenz et al. 2011). Crowds are also susceptible to self-fulfilling prophecies, in which perceived but false popularity can become real over time (Salganik and Watts 2008). Despite these shortcomings, it has been suggested that 'in a well-designed setting, a collective evaluation can match the performance of experts on a given evaluation task' (Riedl et al. 2010). But exactly how wise are the crowds in practical crowdsourcing settings? The question is important for the design of crowdsourcing applications. It would be useful to know to what extent the crowd's judgment can be relied upon and if additional decision-making mechanisms are necessary.

5.4.1 Collection and analysis of statistical data

The quality of the crowd's decisions was analysed by comparing the decisions made by the crowd to the decisions made by the hierarchy. Both OpenIDEO and Threadless epitomise a natural experiment for such a comparison. First, the crowd makes its assessment by viewing, commenting and applauding the concepts or designs. After that the organisations, possibly with the help of experts, decide which concepts to include on the shortlist or what designs to print. For the purposes of analysis, the organisations' decisions are assumed to be correct and the usefulness of the crowd is estimated by looking at how closely their decisions match with the decisions of the experts. Unfortunately, suitable data for this kind of analysis was not available at Quirky.

For OpenIDEO, the Applause phase was selected for further analysis because 1) the availability of data on crowd decisions is better; 2) relying on the crowd in decision making is more useful when there are more options to choose from; and 3) in the Applause phase, the numbers of views and comments and the amount applause are available for each concept in numeric format. In contrast, in the Evaluation phase, the results of the crowd's decision are only displayed in graphical form. Interpreting the crowd's decisions is difficult due to multiple evaluation criteria.

As it is possible for the participants to view, comment on and applaud concepts after the Applause phase is ended, reliable data can only be collected near the time when the OpenIDEO team decides the shortlist. This is why data was collected only on three challenges: How can we manage e-waste and discarded electronics to safeguard human health and protect our environment (e-waste), How might we identify and celebrate businesses that innovate for world benefit – and inspire other companies to do the same (celebrate), and How can we equip young people with the skills, information and opportunities to succeed in the world of work. For the lack of better knowledge on when the team decides the shortlist, data was collected as close to the change of phases as possible, usually a couple of hours before the announcement of the shortlist and once the day after the announcement. In the delayed case, the activity stream on the site showed that only a few comments had been given after the announcement of the shortlist, so the data were not badly contaminated. Descriptive statistics for the collected data are presented in Table 5.15. There is a difference between shortlisted and rejected concepts. The shortlisted concepts have 387 views, 11.4 comments and 11.5 counts of applause on average, whereas the concepts not accepted on the shortlist have 208 views, 5.2 comments and 5.6 counts of applause.

Table 5.12: Descriptive statistics for analysed OpenIDEO challenges

Challenge	**Concepts** (shortlisted)	**Views** Mean (SD)	**Comments** Mean (SD)	**Applause** Mean (SD)
E-waste	106 (20)	233 (181)	7.2 (7.5)	6.8 (6.5)
Celebrate	95 (20)	156 (153)	4.9 (5.2)	5.0 (4.2)
Unemployment	149 (20)	295 (198)	6.6 (6.7)	7.5 (6.5)
Total	350 (60)	239 (190.1)	6.3 (6.6)	6.6 (6.0)

Data on views, comments and applause were normalised to have a mean of 0 and a standard deviation of 1 separately for each challenge to facilitate the comparison of the challenges. The data on the challenges were then combined. Here, we can already see an issue stemming from the use of unrestricted measures for crowd opinion. There is no limit on how many views, comments or the amount of applause that a concept can accrue. The numbers vary from challenge to challenge, which makes it difficult to compare the popularity of concepts across the challenges. This is not only an issue for research but for the crowdsourcing organisation: when aggregated evaluations are relative, a single score does not give much information about the quality of a concept unless it is compared to other concepts within the same challenge. For instance, measuring the applause to view rate could be more helpful in assessing crowd opinion than simple counting.

Due to the natural limit at zero and no limit at the upper end of the scale, the data from OpenIDEO tend to be right-skewed, as seen in Figure 5.2, for the number of views. Distributions of numbers of comments and applause counts have similar shapes. Transforming the data logarithmically made the combined data on views appear more normal. Whereas this is good for the later regression analysis by making the errors more evenly distributed, it complicates the interpretation slightly. The other two variables, the number of comments and the amount of applause, did not benefit much from the transformation, which may cause issues in the regression analysis.

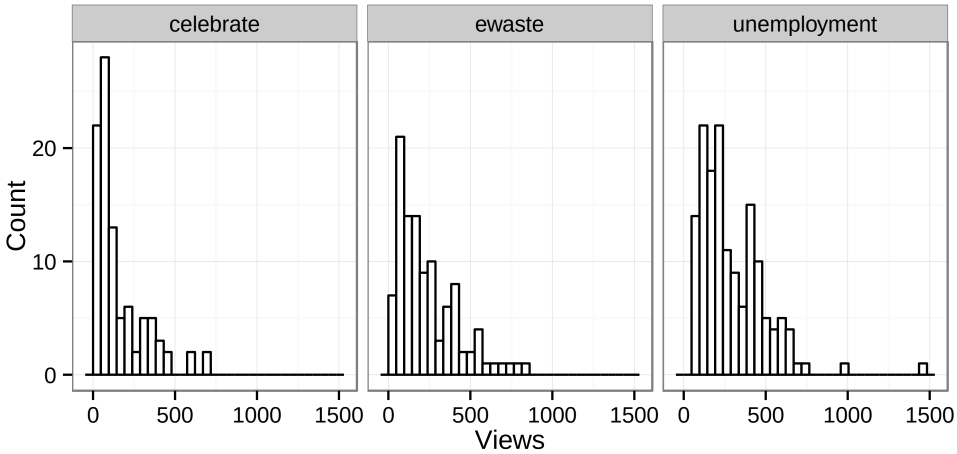


Figure 5.2: Distribution of the number of views for the three OpenIDEO challenges. Distributions of numbers of comments and the applause counts have similar shapes.

Figure 5.3 shows a comparison of shortlisted and rejected concepts from all three challenges. Shortlisted concepts tend to have more views, comments and applause compared to rejected concepts. They also tend to be submitted to the challenge earlier, but this effect is not very pronounced. Outliers at the upper end of the scale are to be expected because the scale is not limited. The outliers are included in the analysis.

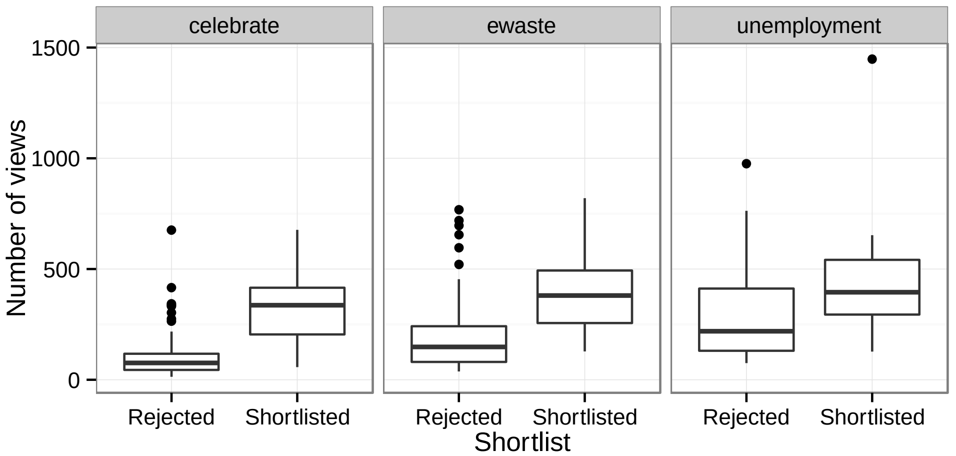


Figure 5.3: Number of views on shortlisted and rejected concepts on three challenges

Threadless describes the selection process of design as follows (Threadless 2013):

> All designs printed on Threadless are voted on and picked by the community. Users can submit designs to Threadless, which are then voted on for a 7-day period by other users in the community. Once the scoring period has ended, the design receives a score from 1 to 5. This is used as a gauge by Threadless to decide what gets made into a tee!

Data from Threadless were collected using a web crawler written in R programming language (R Core Team 2012). The algorithm was slowed down on purpose to prevent disrupting the website during data collection. The dataset contains all designs accepted to the on-going Threadless Challenge (Threadless 2013b) between 8 October 2012 and 1 July 2013. The collected data include the name of the design, user, date of approval, average score of community votes (scale from 1 to 5), number of votes, number of ones, number of fives, and whether the design was eventually printed by Threadless. All this data is publicly available on the Threadless website, although the site does not link printed designs to submitted designs in a consistent way. As a result, multiple approaches were used to assess the status of designs. A design is considered to have been printed if 1) the user who submitted the design has the URL of the submitted design on his or her list of printed designs; 2) the page displaying all designs printed by Threadless (www.threadless.com/all) contains the URL of the submitted design; or 3) a printed design at the Threadless shop has a link to the submitted design. During the observation period, Threadless printed 205 designs out of 15,581 submitted by 12,478 users to the Threadless Challenge. The mean score of all submitted designs was 2.75 (SD 0.39), whereas the printed designs scored 3.18 (SD 0.46) on average. As Figure 5.4 shows, the distribution of average scores resembles quite closely the overlaid normal distribution with the same mean and standard deviation. Distributions of other collected variables, numbers of scores, fives, and ones, are all right-skewed and resemble the distributions of numbers of views at OpenIDEO (Figure 5.2). In this case, logarithmic (base 10) transformation cures the skewness.

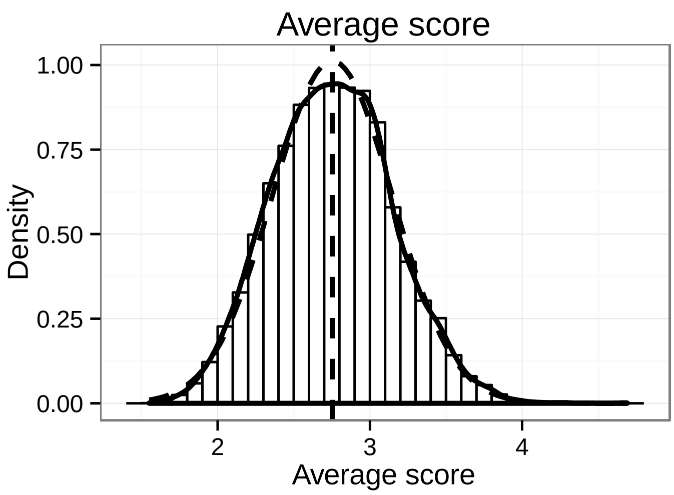


Figure 5.4: Distribution of average scores on Threadless. Normal distribution with same parameters is laid over the distribution in a dashed line.

In Figure 5.5, comparing the variables based on whether the submitted design has been printed on Threadless reveals that there is a tendency for the printed designs to have slightly higher average scores, numbers of scores and number of fives than designs that have not been printed. Numbers of ones tend to be lower for printed designs. However, there is an overlap between all the compared distributions. Many designs with high average scores are not printed, whereas some designs with low scores are. There are five outliers with values less than zero. As the number of actions cannot be negative, these values must have resulted from a processing error either on the website or during data collection. Accordingly, these outliers were removed.

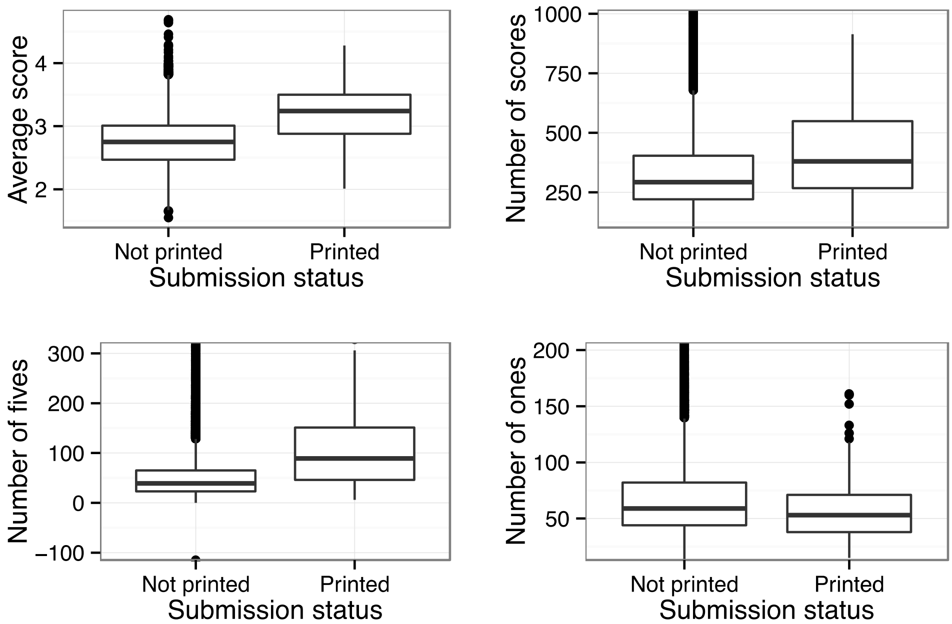


Figure 5.5: Average score, numbers of scores, numbers of fives and numbers of ones versus the submission status on Threadless

The data from both OpenIDEO and Threadless were analysed using the following approach. Seventy percent of the data were randomly selected from the training set and analysed using logistic regression analysis (Field et al. 2012). Data were then implemented in the R environment (R Core Team 2012). Logistic regression was used to model and predict a binary outcome from one or more variables. Given the variable(s), the logistic regression estimates the probability of an observation belonging to one of two possible classes. In the case of OpenIDEO, this means that the model tries to predict the probability of a concept being on the shortlist, given the (normalised) number of views it has gathered. For Threadless, this analysis produced estimates of the probabilities of submitted designs being printed based on the average score given by the crowd. The quality of the models was assessed using the test sets. After that the complete datasets were used to obtain the final results. Detailed analysis reports are presented in Appendices C and D and the anonymised datasets and R scripts[4] for producing them are available at GitHub (Salminen 2015).

[4] An R script is a short program written in R programming language that automates a set of tasks.

5.4.2 Results of statistical analyses

The results of logistic regression analysis on the training sets with odds ratios and their 99% confidence intervals for OpenIDEO are presented in Table 5.13.

Table 5.13: Results of logistic regression analysis for OpenIDEO, odds ratios and their 99% confidence intervals

			99% CI for odds ratio		
	B (SE)	p <	Lower	Odds ratio	Upper
Intercept	-1.85 (0.17)	0.000	0.10	0.16	0.24
Number of views (normalised)	1.02 (0.16)	0.000	1.89	2.76	4.22

The model is statistically highly significant. The probability of obtaining this kind of data without a relationship between the numbers of views a concept has gathered and whether the concept has been selected on the shortlist is essentially zero. The results suggest that when the number of views gathered by a concept in a challenge increases by a standard deviation (relative to the number of views for other concepts in the challenge), the odds of the concept getting to the shortlist increase by a factor of about 3. The relationship between the submission status, the number of comments and the amount of applause, and the order in which the concepts were submitted, were also investigated, but none of these turned out to be statistically significant ($p = 0.22$, $p = 0.20$ and $p = 0.99$) after the number of views was taken into account. There is a moderate correlation between the number of views and comments and the amount of applause (around 0.65), suggesting that a significant part of the information contained in other variables is already present in the number of views. Figure 5.6 displays the normalised number of views for all 350 concepts and their shortlist status. The concepts are presented in their submission order. There appears to be a slight downward trend in shortlisted concepts: the later entries appear to be less likely to be selected on the shortlist. However, this effect was not found to be statistically significant. The figure clearly shows that the attention of the crowd at OpenIDEO, as measured by the number of views, correlates with the decisions of the OpenIDEO staff, but the crowd’s attention does not guarantee that a concept is selected to a shortlist, nor does a lack of attention necessarily mean that a concept it is excluded from one.

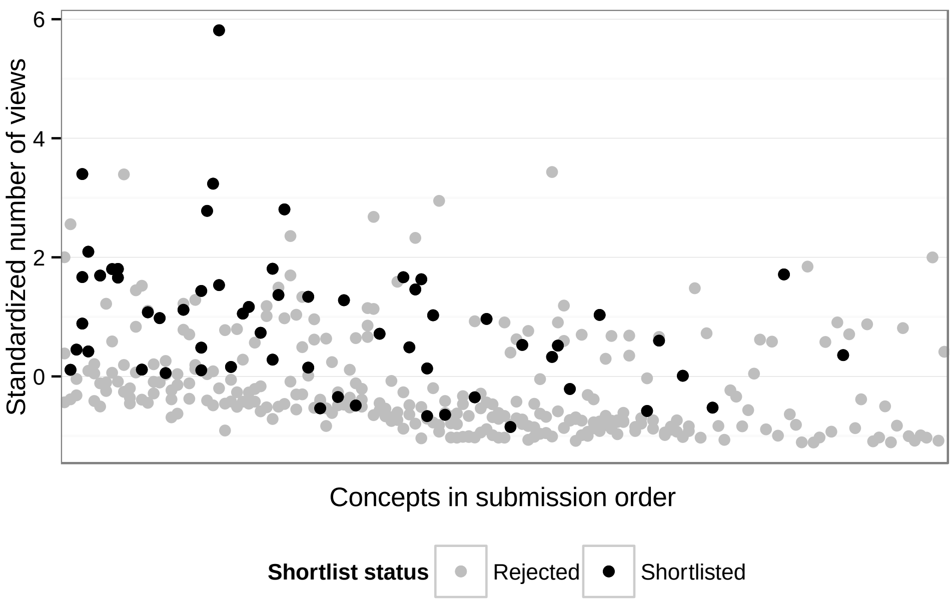


Figure 5.6: All submitted concepts in the OpenIDEO dataset, their view counts and shortlist status. Shortlisted concepts tend to have a high numbers of views, but the difference is not definitive.

The results of the logistic regression analysis on the training sets with odds ratios and their 99% confidence intervals for Threadless are presented in Table 5.14. The comparison of community voting scores with what Threadless actually prints reveals a statistically highly significant effect. When the average score of a design rises by 1 unit on the scale from 1 to 5, the odds of that design getting printed by Threadless increase by an approximate factor of 14. The accuracy of the crowd, however, is not great enough to be relied upon in decision making on its own. The effects of the numbers of scores, numbers of ones and numbers of fives were also analysed, but none were statistically significant ($p = 0.52$, $p = 0.58$, $p = 0.64$).

Table 5.14: Results of logistic regression analysis for Threadless, odds ratios and their 95% confidence intervals

			99% CI for odds ratio		
	B (SE)	p <	Lower	Odds ratio	Upper
Intercept	-12.13 (0.56)	0.000	0.0	0.0	0.0
Average score	2.64 (0.18)	0.000	8.95	14.02	22.15

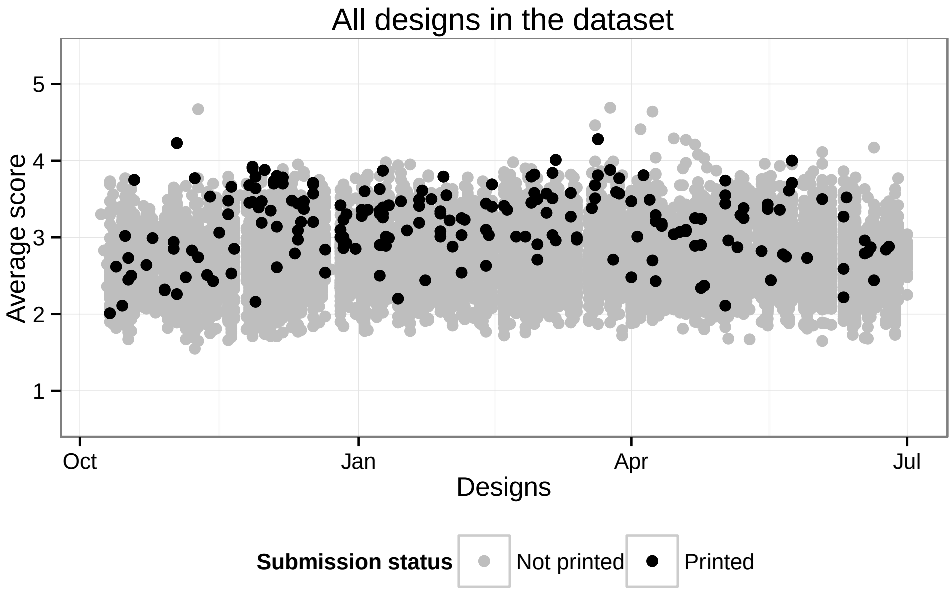


Figure 5.7: All submitted designs in the Threadless dataset, their average scores and submission status. Printed designs tend to have high average scores, but the difference is not definitive.

5.4.3 Validity and reliability of statistical analyses

Particularly in the case of Threadless, the selected designs are small in relation to the number of rejected designs. The models suffer a small penalty when, in rare cases, they predict small probabilities in concepts or designs that were selected. The practical prediction power of the models is weak as they predict that nothing will be selected. The differences between mean squared errors for training and test sets are nevertheless minimal: 0.112 and 0.143 for OpenIDEO and 0.0127 and 0.01264 for Threadless. The models fit the training and test sets equally well, which suggests that they describe the relationship between crowd opinion and the organisations' decisions accurately. The model fit for the Threadless case was also assessed by simulating decision making 10,000 times on the test set and comparing the number of printed designs, their average scores and the standard deviation of average scores to actually observed values. Simulations were also carried out using a baseline model, where each design was given an equal probability of being printed. Simulated decision making with a model from Table 5.14 showed that the observed data are within plausible range in terms of the expected number of printed designs, the average score of selected designs and the standard deviation of scores. The baseline model did not produce average scores matching the observations. More details on the simulation and other model diagnostics are available in Appendices C and D.

5.4.4 Wisdom of crowds in practice

The logistic regression models predict that when the average score of a design or the number of views of a concept increase, the probability of that design or concept being selected also increases. The effects are statistically highly significant, but they are not strong enough to be the sole determinant of what gets selected, as Figure 5.6 and Figure 5.7 clearly demonstrate. These results suggest that the wisdom of crowds may not be high enough in practical contexts to be relied on in decision making, although it may support decision making. Crowds can predict but they cannot categorise. The design of crowdsourcing sites should include secondary decision-making mechanisms in addition to aggregating the crowd's opinions. Violations of the independence condition may decrease the accuracy of the crowd. At OpenIDEO, the statistics for each concept are displayed for participants in real time. Concepts can be sorted by the numbers of views and comments and the amount of applause, among other options. A possible positive feedback loop emerges. A low quality concept that happens to get many views or comments may end up at the top of the list and, due to increased exposure, keep on gathering attention from the crowd. In contrast to the 1 to 5 scale used on Threadless, simple view numbers can easily get inflated as there is no negative feedback loop to balance things out. The crowd cannot dislike a concept. Furthermore, the primary use of numbers of views and comments and applause count may not even be intended to evaluate the quality of concepts. According to an active user at the OpenIDEO user forums, applause, for instance, is supposed to be more about appreciation and positive feedback than evaluation. Although Threadless users cannot see the evaluations of others before voting has ended, the promotion of submitted designs to friends and relatives is encouraged, which may lead to bias in some evaluations. There are also known cases of outright cheating, where some users create multiple accounts to give high scores to their own designs and low scores for everyone else (this is known as downvoting in Threadless community forums). Sometimes, this behaviour can determine the top scoring designs in a challenge. Observations on the website suggest that the accuracy of evaluations is not the only use of community voting. According to the Threadless Forum, many designers consider the average score given by the community to offer important feedback and help their development as graphic designers. Voting on submitted designs can also be a pleasurable activity and thus work as a means of engaging users to spend time on the site, browsing designs that may be on sale in the future.

5.5 Distributed cognition

In the previous section, the crowd's aggregated evaluations were compared to decisions made by the organisation. In a sense, the crowd forms a shared opinion, but how exactly does this happen? The fifth theoretical lens looks at how the crowd interacts during content creation and evaluation. Next, the possible paths of information on each investigated crowdsourcing site are sketched visually. The focus is on the idea generation and evaluation phases of the innovation process. According to the analysis using collective intelligence systems framework summarized in Table 5.11, these phases could manifest collective intelligence.

5.5.1 **Propagation of information at OpenIDEO**

At OpenIDEO, the focus of analysis is on the conceptualisation or Concepting and Applause phases. During these phases, concepts, or solution suggestions to address the problem described in the challenge brief, are created, commented on and applauded. Many events take place at OpenIDEO before the conceptualisation phase begins. OpenIDEO staff and the challenge sponsor first create a challenge brief, which describes the problem at hand but does not give exact specifications on solutions. As the problems are usually difficult and multi-faceted, problem definition has usually continued during the Inspiration phase. Users may have read the brief, commented on it and read other users' comments. Similarly, members of OpenIDEO staff may have read comments and added their own. During the Concepting phase, the challenge brief and associated comments are available on the website so the user can still access them. At the Inspiration phase, users have submitted, viewed, applauded and discussed inspirations. Members of OpenIDEO staff have organised the inspirations into themes in a synthesis meeting. Instructions for the Concepting phase have been posted to the website. All the information created during these preceding steps is available to users during the Concepting phase. Next, the possible pathways of information during concept creation and submission are investigated in more detail.

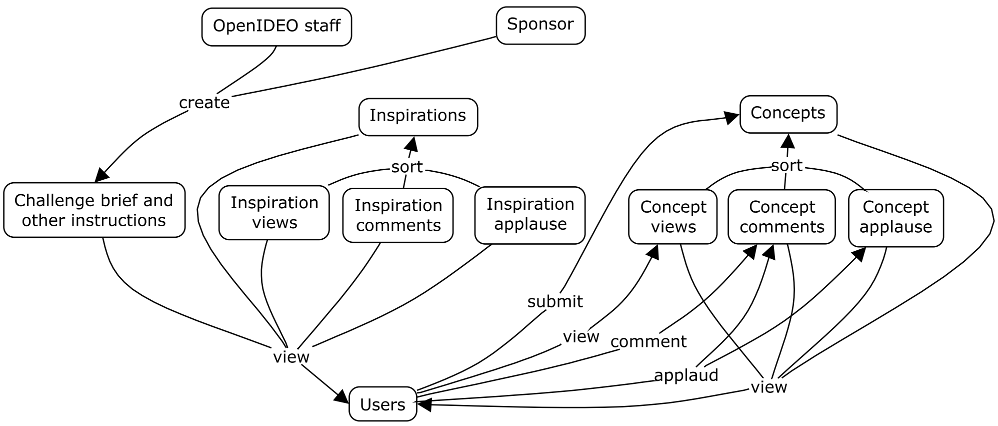


Figure 5.8: Possible pathways of information at OpenIDEO during concept creation

Figure 5.8 depicts possible pathways information can take at OpenIDEO during the Concepting phase. Information originating from OpenIDEO and the challenge sponsor flows through the challenge brief and other instructions to the users, providing guidelines and goals for concept development. The results of cognitive work performed in the Inspiration phase as well as previously-submitted concepts from other users are available and can be used to support the creation of new concepts. Numbers of comments, views and applause counts are available to users at all times. Both inspirations and concepts can be sorted according to these measures. Users thus have access to both absolute and relative popularity measures for the content of the website and can build their concepts upon the earlier contributions of other users. The user is the convergence point of the information pathways and does most information processing. The user must first decode the instructions and wishes of the challenge sponsor as presented in the challenge brief. Then, the user’s task is to select and integrate suitable pieces of information into a concise concept. The OpenIDEO platform provides help in this information-processing task by offering a collection of useful information organised in themes and sortable by different measures of popularity. Finally, the users can comment on and read other’s comments on their submitted concepts. The comments often provide feedback and can help the user to improve the concept before the Applause phase.

In the Applause phase, the challenge brief and inspirations do not play a major role. They can still be viewed, but there is no specific reason why users should look at them. Each concept has a number of views, comments and applause counts associated with it. When users view, comment, or applaud concepts, the corresponding count is increased. Users can see these numbers and sort the concepts according to these popularity measures. Comments can also be applauded. Users can view both comments and their applause count. A lot of information about the current status of concepts is available to users during evaluations. It is easy to see from Figure 5.9 that a feedback loop is formed: user activities increase the numbers of views and comments and the amount of applause. This information is in turn made directly available to users. Users may also sort the concepts according to current popularity rankings, which may have an effect on what concepts they view in the future. Popular concepts may become even more popular just because of the feedback loop.

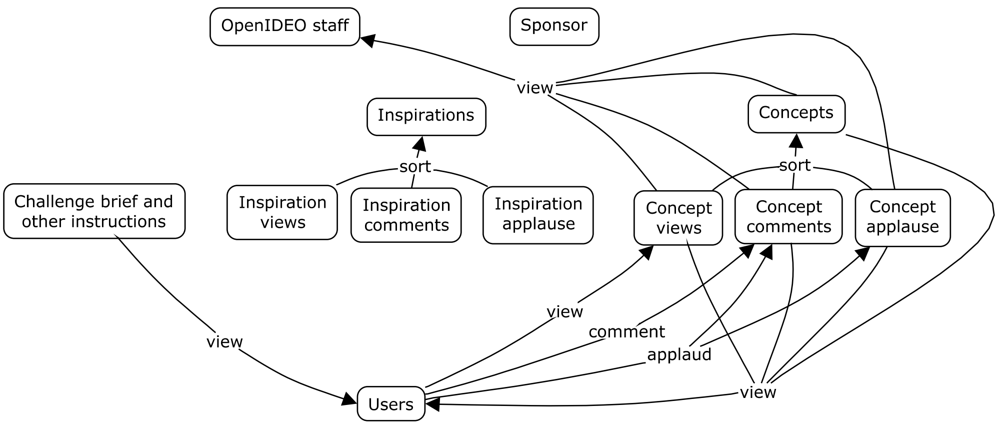


Figure 5.9: Possible pathways of information at OpenIDEO during the Applause phase

After the Applause phase, members of OpenIDEO staff review the results and select about 20 concepts to a shortlist for further refinement. In the Applause phase, information on the popularity of concepts is generated in a system featuring a feedback loop. The information then flows to the OpenIDEO staff for decision making.

5.5.2 Propagation of information at Quirky

In Quirky, the focus of analysis is on the idea submission and community curation phases. There is not much action on the website before users submit ideas. Sometimes Quirky creates specific challenges, but most of the time there are just general instructions on submitting ideas to different product categories at Quirky. Users may view the challenge description and the instructions, and browse and discuss older ideas, but otherwise the possibilities for collaboration on the site before idea submission are limited. Users create and submit their ideas mostly on their own. Figure 5.10 depicts the possible information pathways at Quirky during idea creation and submission. Unlike OpenIDEO, the Quirky platform is not designed to support idea creation before ideas are submitted. As a result, information flows mostly from users to the website. Comments from other users provide feedback about the submitted idea. A user can use this information to update and refine the concept.

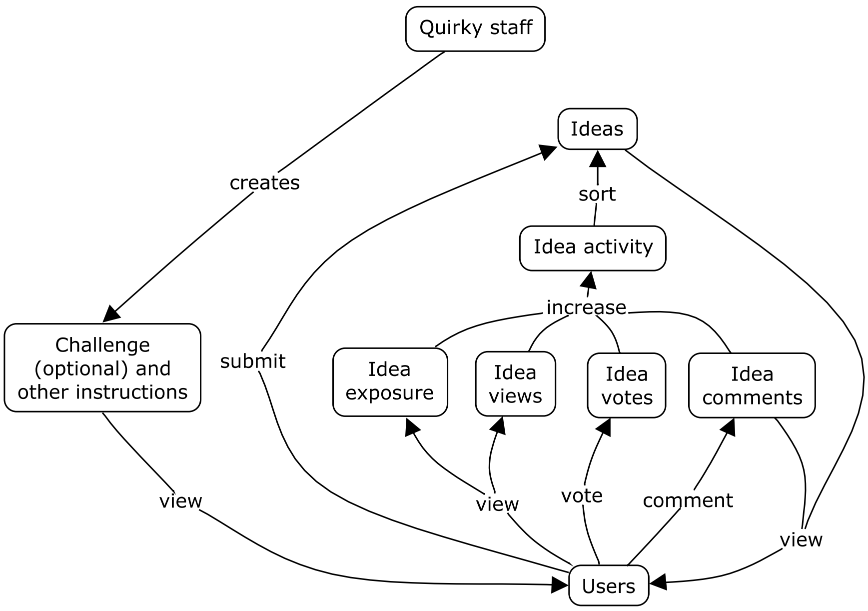


Figure 5.10: Possible information pathways at Quirky during idea submission and refinement

In the community curation phase, the users view, vote and comment on submitted ideas. The amount of information about the other users' opinions is limited. Each idea has a number of exposures, views, votes, and comments associated with it but, apart from comments, users cannot view these statistics directly. The information is only available to Quirky staff. Instead users can sort ideas in evaluation by their relative activity using the Most Active filter, which combines a number of popularity measures. As shown in Figure 5.11, letting the users sort ideas by their relative activity creates a feedback loop. Activities of users increase the relative activity of ideas, which in turn may turn the attention of users towards those ideas.

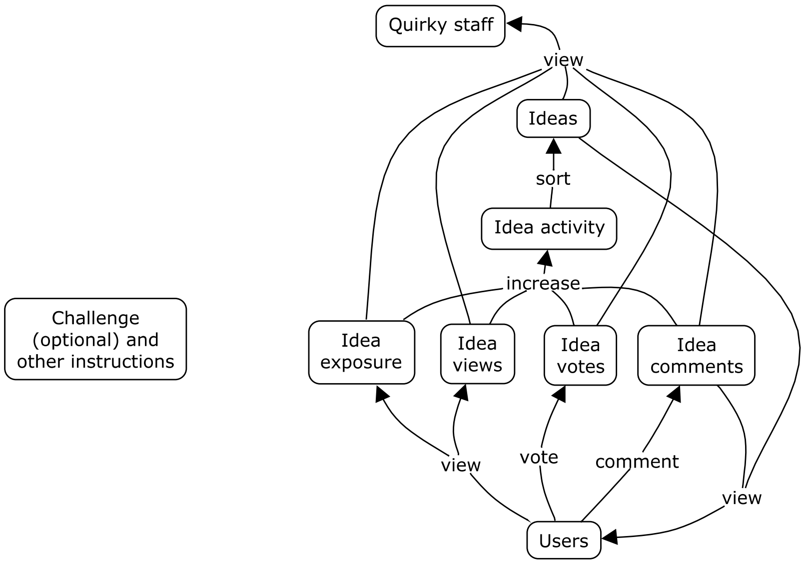


Figure 5.11: Possible pathways of information at Quirky during community curation

After the community curation phase, members of the Quirky staff review all the submitted ideas and select a few of them for further development. As with OpenIDEO, users generate information on the quality and popularity of ideas, interacting in a feedback loop. The information then flows to the Quirky staff for decision making.

5.5.3 **Propagation of information at Threadless**

Analysis of Threadless focuses on the graphic design and scoring phases. Members of the Threadless staff create design challenges, sometimes in collaboration with external partners. The challenge brief is usually a simple, one page description of the theme and possible design constraints. Information from the challenge brief and other instructions are available for users to guide the graphic design process. Users can also view previously submitted designs and their scoring results and comments, which can give them information on what kinds of designs the Threadless community appreciates. Threadless hosts a discussion forum where community members can view and post messages. Some of the discussion threads are collaboration suggestions and playful user-created challenges, which can initiate the creative process. During graphic design, users normally create increasingly detailed sketches of the design they are working on, which effectively transfers information from the designers' minds into artefacts, namely, design sketches. Users can post these sketches to the forum and ask for feedback. They can also view and comment on sketches posted by other users, and read the feedback they have themselves received. Users can learn from the feedback and use it to improve the designs they are developing. When the design is ready, it is submitted to the website. Members of the Threadless staff review all designs before they go up for scoring. Possible information pathways at Threadless during graphic design are displayed in Figure 5.12. Notably there are two feedback loops that can help participants learn and improve their graphic designs: feedback for work in progress from the discussion forum, and scores and comments on finished designs that have already been through the scoring phase.

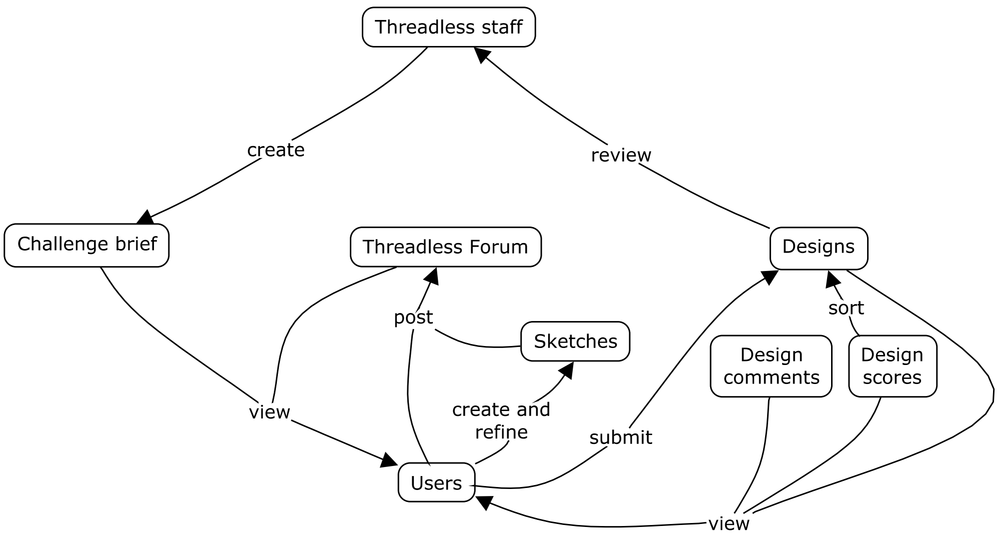


Figure 5.12: Possible information pathways at Threadless during graphic design and submission

After a design has been reviewed and found to be of sufficient quality by Threadless staff, it goes up for scoring. Users view, score and comment on designs. The website keeps track of the average score, the numbers of ones, the numbers of fives, and the total numbers of scores and comments. During scoring, users can view other users' comments but they cannot see information on scores during the evaluation phase. In the scoring phase, all the designs are listed on the Threadless website. They can be sorted by their current average score. As with OpenIDEO and Quirky, a feedback loop is created: popular designs may be viewed more often than less popular ones. The difference is that due to the use of average scores as a measure of popularity instead of, for example, numbers of views, the increased exposure does not necessarily lead to increased popularity, as low scores can decrease the rankings of designs. When the scoring period is finished, members of the Threadless staff select a few designs to be printed. They can use the comments and average scores to help in this decision. Figure 5.13 depicts the possible information pathways at Threadless during the scoring phase. Overall, the picture is similar to OpenIDEO and Quirky: information generated by users in a feedback loop flows towards the decision makers in the organisation.

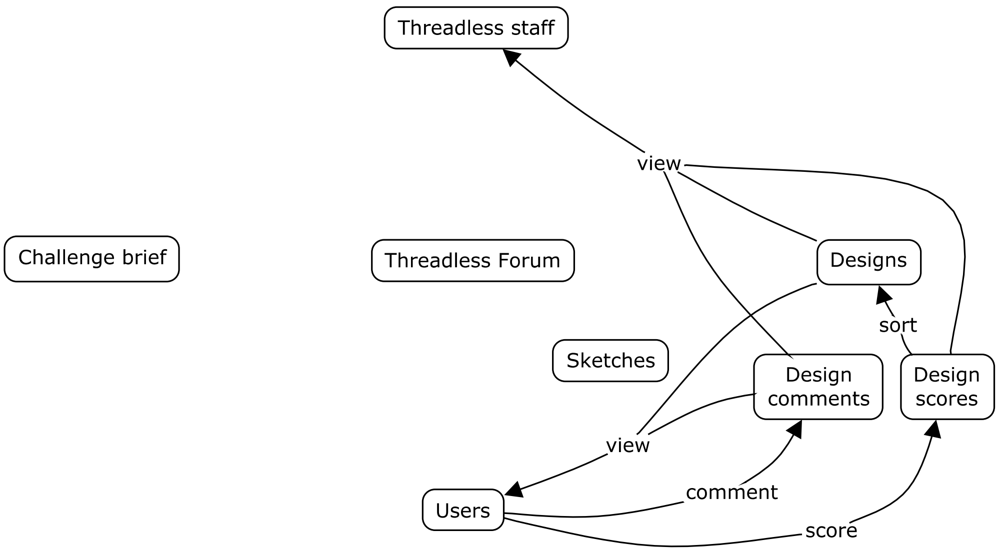


Figure 5.13: Possible information pathways at Threadless during scoring

5.5.4 Feedback loops at crowdsourcing sites

Comparing the flows of information on the crowdsourcing sites two common patterns emerge: 1) the ability to refine submitted content based on the comments of other users and 2) the feedback loop in the evaluation phase. On all three sites, the users submit content to the website in the idea generation phase of the innovation process. Other users can view and comment on the content. The original user can refine and update the content before it is officially evaluated. OpenIDEO offers the most explicit support for this kind of collaboration. Pre-processed information relevant to the task at hand is available in the form of inspirations collected, commented on and applauded by other users. Cognitive work required in the Concepting phase may be reduced, when the user can rely on information someone else has already searched for, evaluated and organised. At Threadless, the discussion forum is used to facilitate collaboration between the users and for soliciting feedback from other users for design work in progress. The feedback helps users to refine their designs and to improve their design skills. On Quirky, there is no pre-submission support for idea generation but, as with the OpenIDEO platform, the submitted ideas can be refined and updated during evaluation, possibly taking advantage of the feedback from other users in the comments. Figure 5.14 presents a generalisation of these feedback loops. Users submit content to the website. Other users can view and comment on the submitted content. The original user can view the comments, respond to them and use them to refine the submitted content. The crowdsourcing platforms allow information to flow between participants; different individuals can process it. The individual user submitting the content is still a focal point of action as he or she is taking care of most of the information processing, for example, by making a t-shirt design or writing down a product idea.

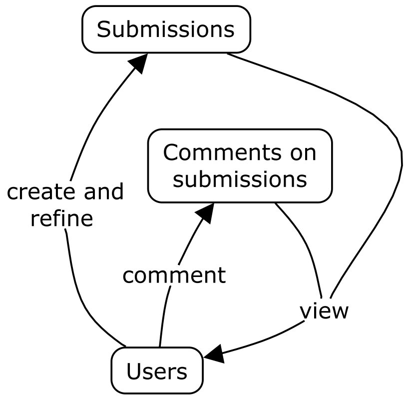


Figure 5.14: Feedback loops taking place during content creation, submission and refinement

A feedback loop during the evaluation phase can be found in all the investigated crowdsourcing systems. The feedback loop is shown in Figure 5.15.

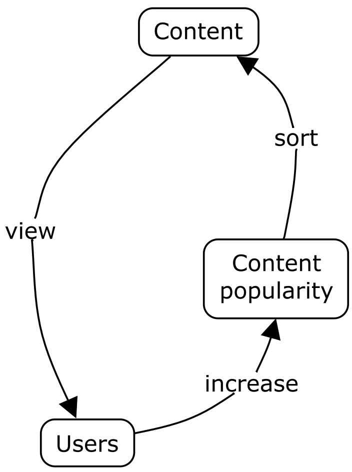


Figure 5.15: Evaluation feedback loop found in all investigated cases

At OpenIDEO, the popularity of a concept, measured in terms of views, comments and the applause it has garnered can increase its visibility because users can sort the concept by these measures. Increased visibility may then increase the probability that users view, comment or applaud the concept, which in turn increases the popularity of the concept. At Quirky, users have an option to view a list of the Most Active ideas. Users notice the ideas near the top of the list most easily, which can lead to even more activity on those ideas, boosting their position on the Most Active ideas list he even further. At Threadless, users can sort the designs by their current average score. As with OpenIDEO and Quirky, a high average score can increase the likelihood that the user scores a design. The difference is that at Threadless increased scoring activity can both increase and decrease the design's average score, depending on the user's opinions of the design. Threadless is the only one of the three sites with the possibility of both positive and negative feedback loops during idea evaluation. At the other two sites, a concept or idea can drop down the listing only when some other idea gets even more attention. The existence of feedback loops during idea evaluation could have consequences on the wisdom of crowds and could partially explain the differences between the crowd's opinion and expert decisions. Feedback loops create dependencies between user evaluations, which violate the independence condition necessary for a crowd to make accurate assessments. The question arises as to how important these feedback loops are for the accuracy of the crowd and whether there differences between the effects depend on the choice of mechanism for aggregating crowd opinion.

5.5.5 Simulation model

As understanding feedback loops and nonlinearities is challenging even in very simple systems, a simple simulation model is built to explore the effects of identified feedback loops on the accuracy of crowd evaluations. A common feature in all the cases is that during the evaluation of ideas, concepts or designs, the content is displayed to the users by default in random order with regard to the current popularity of the item. However, users have an option to sort the content by one or more popularity measures, such as average score, number of views or votes, or idea activity. This leads to the possibility of feedback loops on the sites: popular content may get increased attention, which in turn can increase popularity even further. To investigate the role of the feedback loops, a simple agent-based simulation model was constructed using AnyLogic simulation software (AnyLogic 2015; Borshchev 2013).

Agent-based modelling is the most recently developed approach to modelling. The advantage of the approach is that models can be built bottom-up based on knowledge about the behaviour of individual system components (Borshchev 2013). At the moment, standardised languages for agent-based modelling do not exist. Instead, models are usually created using a combination of graphical editors and scripts. Here, the model is built according to the Threadless scoring process. A few additions were made to enable comparison with other cases. The purpose of the simulation model is to 1) estimate the effect these feedback loops have on the accuracy of the crowd's evaluations and 2) compare the accuracy of different evaluation approaches, namely, the averaging on a scale from 1 to 5 as used on Threadless and vote count as used on OpenIDEO and Quirky. The model simulates the interactions of two entities: users and designs. Designs are assumed to have an inherent and objective true quality value, which is measured on a scale from 1 to 5. Users visit the site, evaluate designs either in random order or sorted by their current popularity, as measured either by average score or vote count, and leave the site after getting bored. The simulation model can interpret the evaluations both as scores on a scale from 1 to 5 and votes. In the end, the aggregated evaluations of the users are compared to the true quality scores of the designs. The strength of the feedback loop (the probability that users sort the designs) can be varied to investigate the role feedback plays on the accuracy of evaluations. Figure 5.16 displays a state chart from the model that describes the simulation logic. The simulation model is next described in more detail.

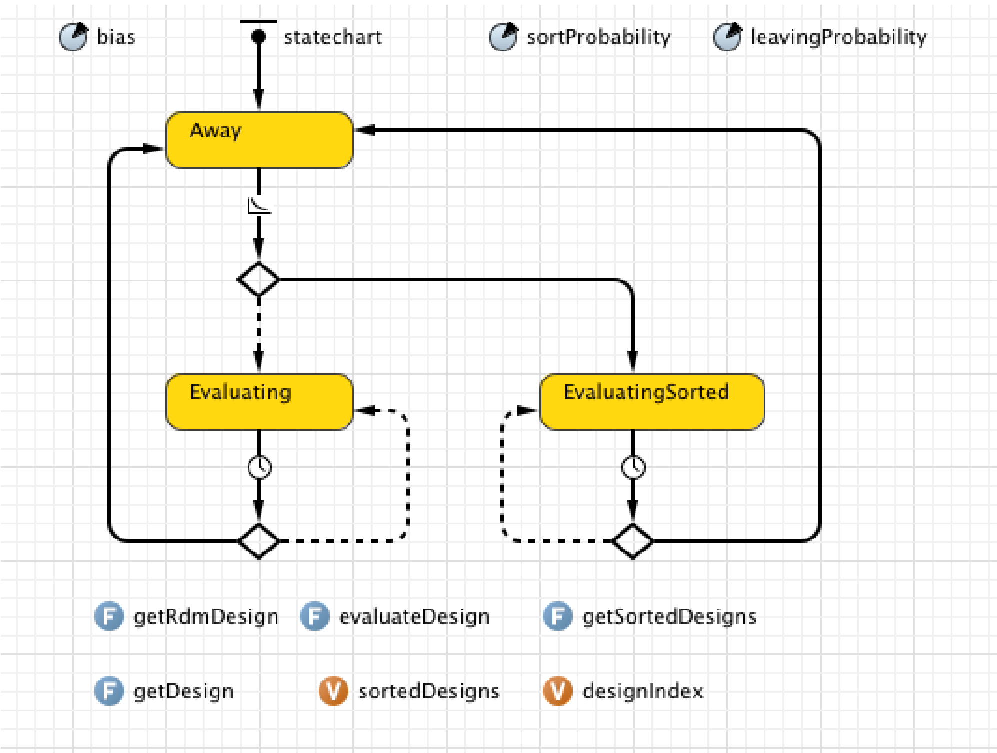


Figure 5.16: Simulation state chart from AnyLogic simulation model describing the logic of the agent-based simulation

Users. Users are modelled as agents. Each user has an individual bias, a tendency to systematically overestimate or underestimate the quality of designs. The bias is determined by a draw from triangular distribution (min -1, max 1, mean 0). Except for this individual bias, the users are identical. They visit the site at an average rate of three times a day, which is sufficient to produce a realistic total number of evaluations. After arriving, the users start evaluating designs either in random order or sorted by their current average score or vote count. Which option the user chooses is determined probabilistically, and the probability of users choosing the sorted order can be varied to adjust the strength of the feedback loop. The higher the probability, the more likely the users are to evaluate the currently most popular designs. Users keep evaluating designs until they get bored and leave the site, to return again as determined by the visiting rate. After each evaluation, the user gets bored with a given probability and leaves the site.

Designs. Each design has its own true quality, which is measured on a scale from 1 to 5. Statistical data collected from Threadless are used as a basis for the true quality scores. Another option would have been to generate the scores by drawing them from some suitable distribution. However, using real data automatically provides a realistic distribution of quality scores and allows for tuning the simulation to use a similar number of evaluations. The actual values of quality scores are not important for the results as the goal is to assess how closely the crowd can match them. The designs keep count of the number of times they have been evaluated, the number of different scores given, the current average score, and the number of votes.

Evaluation. When a user evaluates a design, the following actions take place.

1. The design is evaluated by adding together the true quality of the design, the user's individual bias term, and a randomly-generated evaluation error. The result is rounded to the closest whole number between 1 and 5.
2. The evaluated design's score count for the corresponding score value and view count are increased by one.
3. If the user's score is higher than a threshold value, the vote count of the design is increased by one. This reflects the assumption that users will vote for designs they like, and do nothing otherwise.
4. An evaluation is assumed to take 30 seconds (in simulation time), after which the user takes the next action.

Parameters and tuning. During a simulation run, 100 different users evaluate 1,000 randomly selected designs from the Threadless dataset. One agent may evaluate the same design several times. This is different from the websites where an individual user can vote on or score the same piece of content only once. For the purposes of the analysis this difference is not important. The agents in the simulation model only represent stereotypical users and as the only difference between the individual agents is their bias term, the number of meaningfully different agents is limited. A number of parameters about the model can be adjusted to produce different simulation results. These parameters, their effects on the simulation and selected values are summarised in Table 5.15. The evaluation period was chosen to be seven days, which is the same period that Threadless uses. The visiting rate and probability of the user leaving were then adjusted to reach a similar total number of evaluations as observed on Threadless. In the first set of simulation runs, the probability of sorting varied to explore the effect of feedback loop strength on the accuracy of crowd evaluations. In the second set of simulation runs, the probability of sorting was adjusted between 0 and 0.75 and the evaluation time was varied to understand the dynamics of the system in the long run. Results of the simulation runs with different values of sorting probability were saved as comma-separated files and analysed visually using the R programming environment. The accuracy of crowd evaluations was estimated by calculating the Pearson correlation between true quality score and simulated average score, and true quality score and vote counts.

Table 5.15: Simulation parameters and their justification

Parameter	Effects	Selected value(s)	Justification
Number of users	Number of evaluations Number of different users	100	Covers the range of possible bias values beyond meaningful differences. Chances are that several agents are practically identical because evaluations are rounded
Bias	Systematic error in individual user's evaluations	Triangular distribution (-1, 1, 0)	Ease of use. Purpose is just to add some random noise to the evaluations.
Visiting rate	Number of evaluations	3 per day	Total number of evaluations and their distribution between designs, similar to Threadless data
Probability of user leaving	Number of evaluations Distribution of number of evaluations when sorted designs are evaluated	0.005	A low value was selected to decrease the effect of sorting designs. Even with this low value, the effect is very noticeable.
Voting threshold	Number of votes Accuracy of evaluations	3	Users are assumed to vote for everything they consider better than average. This value provides the best accuracy when sorting probability is 0.
Simulation time in days	Number of evaluations	7	Same as on Threadless
Probability of user sorting the designs	Distribution of evaluations between the designs	0–1 in 0.1 increments	Used to explore the effect of feedback loop strength on evaluations

The simulation model reveals that the feedback loop alone can damage the accuracy of the crowd and that there is an important difference between the dynamics of two evaluation schemes used at crowdsourcing sites. A feedback loop in the crowdsourcing system can decrease the crowd's overall accuracy when vote counts are used as an aggregation method. The feedback loop directs the crowd's attention to a subset of content, and the evaluation of other designs is ignored to an extent. As a result, some part of the content is evaluated more thoroughly than others. In the absence of feedback loops, both average scores and vote counts are highly correlated with the true quality of designs but, as the strength of the feedback loop increases, accuracy suffers, as shown in Figure 5.17. Average scores do better in comparison to voting, but the presence of a strong feedback loop can still slightly slow down the convergence of crowd evaluations to correct values. When the quality of designs is estimated by vote counts, the feedback loop permanently damages the crowd's accuracy. The accuracy of the crowd increases only slightly over the long run. The difference to accuracy of average scores is dramatic, as can be seen in Figure 5.18.

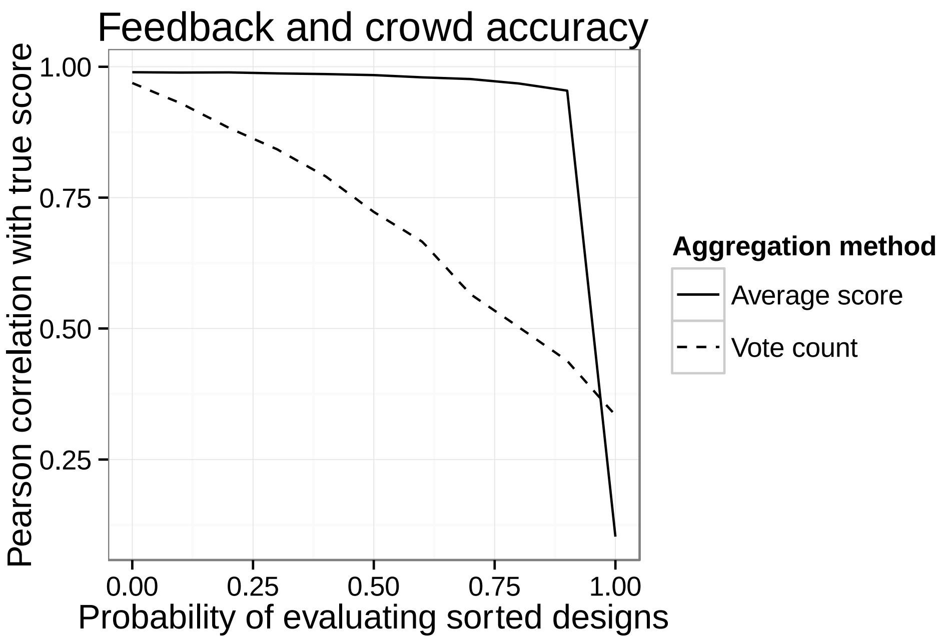


Figure 5.17: Effect of feedback on crowd accuracy

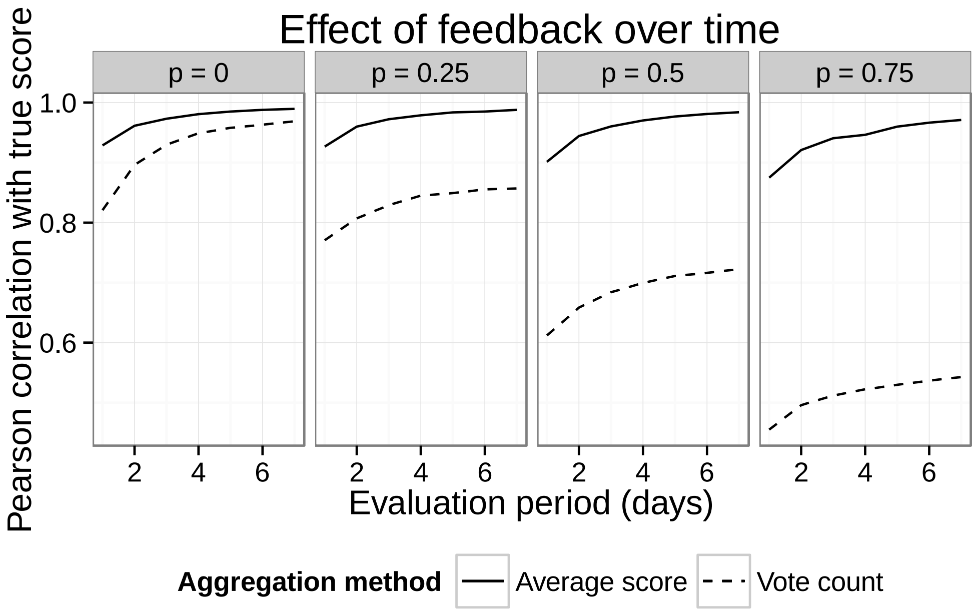


Figure 5.18: Effect of feedback over long evaluation times

When users cannot sort the designs, both average score and vote count closely reflect the true quality. When users sort the designs most of the time (sorting probability = 0.75), according to the current crowd estimate, the average score will still converge to the correct value, but more slowly, while the vote count will never become accurate.

6 Discussion

Crowdsourcing in the context of innovations typically involves a crowd creating a collection of mostly independent contributions, usually in the form of ideas or concepts. The general framework of innovations consisting of problems and solutions appears a good fit with the actions that take place at the investigated crowdsourcing sites. This should not be a surprise as one of the definitions of a crowdsourcing system is a system that enlists a crowd of humans to help solve a problem defined by the system owners (Doan et al. 2011). This thesis set out to study whether something similar to the swarm intelligence of social insects might be going on at crowdsourcing sites that develop innovations.

First, the concept of collective intelligence in humans was clarified by conducting a systematic literature review and building a theoretical framework. Then, three cases of crowdsourcing innovations were investigated in detail mostly using the ethnographic research approach. In all the investigated cases, crowdsourcing was used to generate solution suggestions. On Threadless, the approach appeared to be working well, perhaps because the problem of creating funny, cool, original or otherwise interesting t-shirt designs is easily understood by the participants. In contrast, at OpenIDEO and Quirky, the problem to be solved is not as clearly defined. OpenIDEO focuses on difficult problems. Understanding the issue is therefore an important part of the challenge. At Quirky the problem definition is almost fully at the hands of the participants. The idea submission form requires the user to explain what problem the submitted idea is supposed to solve. Scrolling through the submitted ideas gives an impression that Quirky might be suffering from poor quality problems. This observation was corroborated by a streamed feedback show on the site where Quirky staff discussed the importance of the problem (Quirky 2013). An approach that focused more on problem sourcing suggested by Cummings et al. (2012) might be helpful. In problem sourcing, the crowd is used to identify interesting problems for a product development organisation to solve. In their study, the approach was found to work well and to provide several benefits to the organisation (Cummings et al. 2012). However, here the main focus is on the possible roles collective intelligence plays in crowdsourcing innovations.

Table 6.1 lists the main findings and their implications using different theoretical lenses in cross-case analyses. Different viewpoints on collective intelligence complement and extend each other, drilling down into the details. The collective intelligence genome describes the cases using a common framework, highlighting the differences and similarities in the processes. Below, the identified similarities are used to describe the analytical case to which the results can be generalised. Interpreting the creation of innovations as search problems enabled a comparison of crowdsourced innovation processes to the search process of individual inventors. It appears that the crowdsourcing sites can better be described as collectors of individually created inventions rather than collective inventors. The users do most of the cognitive work on their own rather than being integral parts of a larger cognitive system. Assessing the cases in the framework of collective intelligence systems, corroborates the previous conclusion. All the enabling properties of collective intelligence systems are easily observed, and the emergence of being ‘more than the sum of the parts’ between local and global levels of the system becomes the defining factor for collective intelligence. Two possible manifestations of collective intelligence were identified: forum-like discussions between the participants and crowd evaluation. The second of these manifestations, the wisdom of crowds effect in the evaluation of crowd submissions, was assessed statistically by comparing the crowd evaluations to expert decisions. Relying on the theory of distributed cognition, possible pathways of information were then tracked to better understand interactions taking place on the crowdsourcing sites. In each case, a feedback loop was found, which could influence the crowd’s behaviour during the evaluation. A simple simulation model was built to investigate the effects of feedback on crowd accuracy. Feedback loops can reduce the accuracy when vote count is used to aggregate the crowd’s opinion.

Table 6.1: Main results from cross-case analyses.

Theoretical lens	Findings	Implications
Collective intelligence genome	Work between crowd and organisation distributed differently in each site	Similarities between the cases define the analytical case, for which the generalisation is valid
Innovation as a search problem	Each individual needs to go through the phases of inventor’s search process	Collection of individuals rather than a collective mind
Collective intelligence systems	All enabling properties are present. Emergence the main determinant for collective intelligence. Found in forum discussion and evaluation.	Forum discussion and crowd evaluation potential manifestations of collective intelligence
Wisdom of crowds	Correlation between crowd evaluation and expert decisions	Crowd can predict but cannot categorise
Distributed cognition	All sites have a potential feedback loop in evaluation	Feedback loop could influence system behaviour
Simulation model	Feedback loop can reduce the accuracy of the crowd when vote count is used for aggregation	Averaging is a preferable aggregation mechanism in the presence of feedback loops if accuracy of the crowd is important

6.1 Analytical case

The three cases can be summarised based on their common features to form a generalised analogical case of crowdsourcing innovations. The summary is presented in Table 6.2 in the form of the collective intelligence genome. The findings of this study should be valid for crowdsourcing sites that match the genome. The generalised innovation process starts with a challenge published by the crowdsourcing organisation. A crowd comes up with suggestions and submits them to the organisation, creating a collection of content. The suggestions are open for anyone to view and discuss, and may be refined and improved based on the feedback from these discussions. The crowd evaluates the suggestions using some form of voting scheme. The voting schemes allow the emergence of feedback loops, which may decrease the accuracy of evaluations if simple counting is used to aggregate the results. In addition to supporting decision making, crowd evaluations may have other purposes and uses for which the feedback loop may be beneficial. After crowd evaluation, the crowdsourcing organisation selects winning suggestions for further development. The selection may partially rely on crowd evaluations and comments, but the crowd does not make the final decisions.

Table 6.2: Collective intelligence genome of a generalised analogical case. The results of this study should be valid in other cases with similar genomes.

Process phase	**What**		**Who**	**Why**	**How**
Challenge	Create	Innovation challenge	Hierarchy	Extrinsic	Hierarchy
Idea generation	Create	Solution suggestions	Crowd	Intrinsic or extrinsic	Collection
Feedback	Decide	How to improve the solution suggestion	Crowd	Intrinsic or extrinsic	Consensus
Refinement	Create	Improved solution suggestion	Crowd	Intrinsic or extrinsic	(Collaboration)
Evaluation	Decide	Average score or vote count	Crowd	Intrinsic or extrinsic	Averaging or voting
Winners	Decide	Winning solutions	Hierarchy	Extrinsic	Hierarchy
Development and implementation	Create	Final product	Hierarchy	Extrinsic	Hierarchy

6.2 Theoretical framework for collective intelligence

A sample of literature discussing collective intelligence in humans was reviewed and the themes were categorised into micro-, macro- and emergence-level phenomena. The framework is similar to the conceptual model of Luo et al. (2009), the gist of which is the question of how macro-level phenomena emerge from micro-level interactions. The proposed framework emerged from data collected from contemporary literature. Therefore, it is arguable that the scientific community has already implicitly divided collective intelligence into these three levels of abstraction. Making this division explicit hopefully brings some structure to the discussion and helps fit the pieces of the puzzle together. The categorisation of themes related to collective intelligence (Table 3.3 and Figure 3.4) provides guidance for selecting topics for further literature reviews and suggests how the results might fit into the big picture of collective intelligence in humans.

The framework allows some educated guesses on how other phenomena might be connected to collective intelligence. For instance, promising results have been obtained from using theatre-based methods in relieving organisational issues (Pässilä and Oikarinen 2011). As ‘improvisation theatre is about interaction’ (Partanen 2011), it can be hypothesised that theatre-based methods contribute to collective intelligence by influencing human interaction at the micro level. Visualisation tools for group work, such as sticky notes and shared visual templates (e.g., Sibbet 2010), could be interpreted as shared, dynamic memory systems, which facilitate the functioning of complex adaptive systems. Finally, using Twitter searches to monitor discussions on social media arguably increases the probability of stumbling upon some relevant new information. Complex interactions of millions of users manifest themselves as a probabilistic phenomenon in a way that has even been compared to the workings of a brain (Pomerlau 2009).

6.3 Forum discussion as a form of collective intelligence

Next, the finding of forum discussions as a possible manifestation or form of collective intelligence is applied to the framework of collective intelligence derived from the literature. Figure 6.1 extends Figure 5.14, placing it in the larger context of the theoretical framework of collective intelligence. Content from Figure 5.14 is presented in continuous lines, whereas the rest of the framework (Figure 3.4) derived from literature is in shown in dashed lines. Here creation and refinement of content and commenting have been collapsed in a single arrow for clarity.

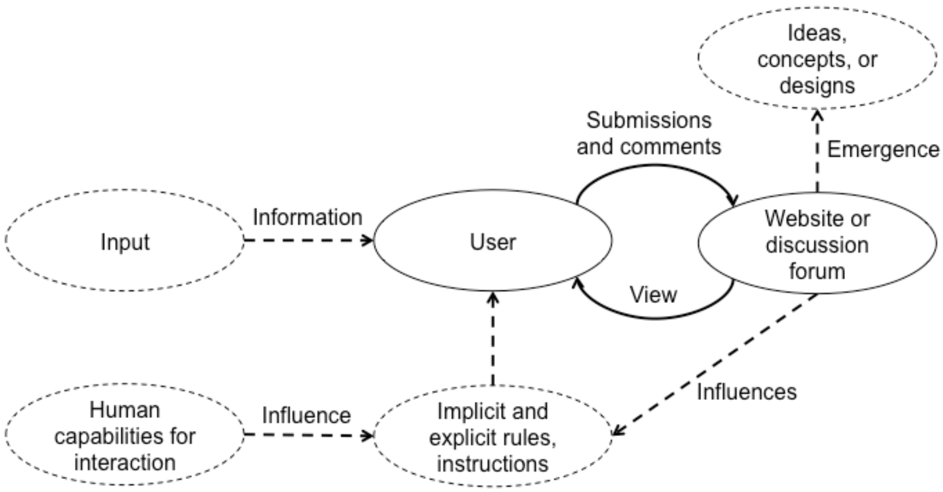


Figure 6.1: Discussion as a form of collective intelligence. Content from Figure 5.14 is presented in continuous lines, whereas the rest of the framework (Figure 3.4) derived from literature is in shown in dashed lines.

Users bring new information into the system from the environment and synthesise it in submissions, comments and evaluations. Users do most of the information processing. They can change the state of the distributed memory by making submissions and writing comments. Feedback to users comes in the form of comments from other users. Interaction rules consist of the implicit rules of the particular cultural context and the constraints and instructions of the website, such as possible character limits for messages. The website functions as a distributed memory that facilitates the discussion of participants. A consensus may emerge as a result of discussions. Interaction resembles normal human group behaviour, details of which are already well known. For a thorough review on human group performance see Kerr and Tindale (2004), and Conradt and List (2009) for a comparison to other animals. These reviews give a good overview of the vast amount of research on the topic. Diverging from the mainstream cognitive sciences, the assumption of individuals as the correct unit of analysis has recently been challenged (Gallagher 2013; Merrit et al. 2013; Pentland 2007). The importance of the social networks containing the individuals has been pointed out: in certain situations, the network effects can predict 40% or more of human behaviour (Pentland 2007). Goldstone and Gureckis (2009) review several case studies on collective behaviour where they consider four attributes: the motivation of individuals, interactions and typical dynamics resulting from the interactions, and characteristic outcomes at the group level. The results of case studies are compared to individuals, swarms, innovation diffusion and online communities, which leads to a discussion on the reducibility of group behaviour to individual behaviour and to a question about the possibility of group cognition (Goldstone and Gureckis 2009). The collective intelligence factor identified in large-scale experiments (Woolley et al. 2010) might play a role in forum discussions found on the crowdsourcing websites, although controversy on the factor still exists. The collective intelligence factor might be a manifestation of the general factor of personality at the group level (Woodley and Bell 2011). In addition, a recent study on the collective intelligence factor failed to find it in a virtual context (Barlow and Dennis 2014). In the context of idea generation, group work and especially building on the ideas of others appears to increase the quality of ideas, at least in an experimental setting (Kohn et al. 2011).

6.4 Wisdom of crowds as a form of collective intelligence

Below in Figure 6.2, the second possible manifestation of collective intelligence, the wisdom of crowds in the evaluation of submissions, is interpreted in the framework of collective intelligence derived from the literature. The figure extends Figure 5.15 and represents it in the context of the framework for collective intelligence.

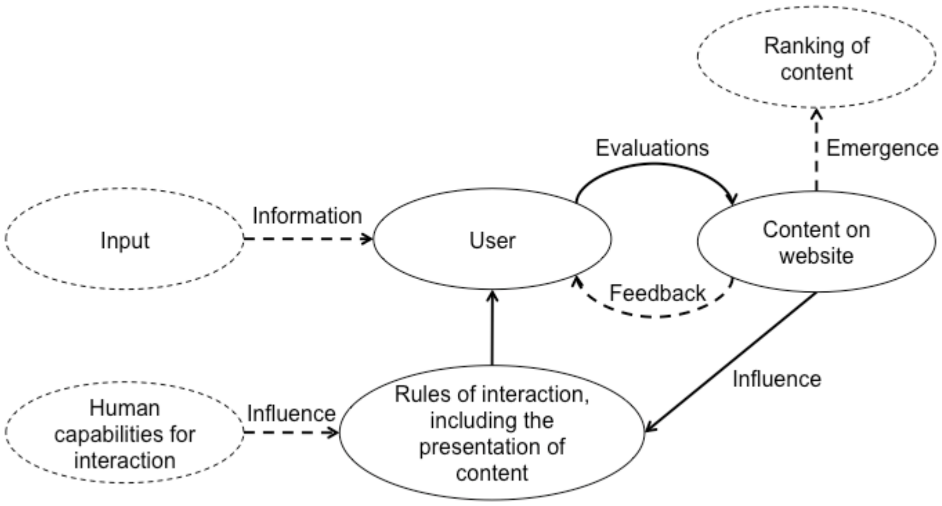


Figure 6.2: Evaluation of content as a form of collective intelligence. Content from Figure 5.15 is presented in continuous lines, whereas the rest of the framework (Figure 3.4) derived from literature is presented in dashed lines.

Users evaluate content based on the information they bring into the system from the environment. The activities that users can take to alter the state of the distributed memory consist of viewing, commenting and evaluating content. The earlier evaluations, views and comments are aggregated and turned to feedback to new participants: popular content may be more likely to be viewed because the content can be sorted based on the current rankings. Rules are again cultural and website-defined but here the characteristics of the website are perhaps more important. The most important rules are concerned with the format of evaluation, and how and when the users participate. Distributed memory keeps track of the scores and votes for each piece of content. Order emerges as a result of the crowd interacting with the website: the ranking of content reflects the aggregated opinion of the crowd to a degree. In contrast to the previous figure, the feedback loop takes a different route. User evaluations influence the ranking of content on the website, which changes the rules dictating what content is shown to the users.

When it comes to the wisdom of crowds, both simulation models (Hong and Page 2004) and real-life experiments (Krause et al. 2011) have demonstrated that under certain conditions groups of low-skilled problem solvers can outperform a group of high-skilled problem solvers. A group can be more intelligent than the individuals making up the group. The potential benefits from this kind of collective intelligence depend on the type of the problem to be solved. If the problem is suitable, increasing the diversity of problem solvers may be more beneficial than increasing their expertise. However, diversity alone is not enough to ensure intelligent performance at the group level. Humans are very adaptive beings and even minor social influence can undermine the wisdom of crowds, as a study by Lorenz et al. (2011) shows. They identified three detrimental effects for the wisdom of crowds: the social influence effect, the range reduction effect and the confidence effect. In line with the previous research, this study also found that social influence in the three cases it investigated was able to undermine the wisdom of crowds in practical contexts. In addition to the three effects already identified, indirect social interaction may undermine the wisdom of crowds in another way. The feedback loop that increases the visibility of the most popular content on the site can delay the convergence of crowd opinion or prevent it altogether depending on the aggregation method used: the crowd focuses on the most popular content and ignores the rest. The results of agent-based simulations closely resemble the findings from experiments with artificial cultural markets, where the researchers were able to artificially invert the true popularity of songs in the web-based music market by reversing the original popularity ratings for the songs. Although, the artificial distortion reduced the correlation between approval ratings and popularity, in the long run, the best songs were able to recover popularity (Salganik and Watts 2008; 2009).

In addition to evaluation accuracy, there may be other important factors to consider when designing a crowdsourcing system. For instance, voting can work as a form of marketing when participants spend time on the website evaluating the content. Voting results give significant feedback to the participants. These two uses may place different demands on the voting mechanisms. As users may prefer to evaluate high-quality content, and to increase the amount of time they spend on the site, crowdsourcing organisations may offer them the chance to view the content currently deemed the most interesting by the crowd, sacrificing some evaluation accuracy. This speculation is in line with the findings from a study where idea competitions were put under scrutiny. Mortara et al. (2013) found that although ideas from such competitions are often difficult to put in use, these competitions provide other important benefits for the entities organising them. Idea competitions are a way to gain improved intelligence on general opinions and to improve public relations through closer interaction.

6.5 Validity and reliability

Common criteria for judging research quality are construct validity, internal validity, external validity and reliability. Construct validity means that correct operational measures are identified for the concepts being studied: the study investigates what it claims to investigate (Gibbert et al. 2008). The phenomenon of interest should be defined in terms of specific concepts. If possible, operational measures matching those concepts should be identified with citations to already published studies using the concepts in similar ways. Construct validity is important especially during data collection. According to Yin (2008), using multiple sources of evidence and establishing a chain of evidence helps achieve construct validity. Multiple sources of evidence allow triangulation in terms of data, researchers, methods and even perspectives, if several theoretical lenses are used. Arriving at the same results from different starting points makes their correctness more believable and mitigates any bias present in the data sources, researchers and analytical methods. An intact chain of evidence makes it possible to follow the analytical chain backwards from the conclusions all the way to the raw data. There should be no logical gaps along the way.

Internal validity is concerned with establishing a causal relationship between the variables of interest and distinguishing the actual relationships from spurious ones (Yin 2008). It is about inferences, or explaining how one thing leads to another. An inference is made every time an event cannot be observed directly. The question is, is the inference correct, and have the alternative explanations been taken into account? Internal validity arises during data analysis in studies that seek to explain causes. In descriptive or exploratory studies, internal validity is not an issue (Yin 2008). A clear research framework, the comparison of observed patterns with predicted or earlier findings, and theory triangulation, all help in establishing internal validity (Gibbert et al. 2008).

External validity or generalisability refers to defining the domain where research findings can be generalised. The conclusions of a study must be shown to be valid in other contexts and not only in the setting where they were originally developed (Gibbert et al. 2008). A typical approach in survey research is to use statistical generalisation, where results obtained from a statistical sample are generalised to the population as a whole. In case studies, analogical generalisation is more common, where a particular set of results is generalised to some broader theory. Similar results can then be expected from other cases where the same theory applies. External validity arises when dealing with theoretical frameworks and case selection. Replication logic and correct use of theory are useful tactics for increasing the external validity of a case study. If the results of one case study can be replicated in another case, the argument for the generality of the conclusions becomes stronger (Yin 2008).

Reliability is about demonstrating that the study is free of random errors. It should at least in principle be possible to repeat the actions taken during the study, such as data collection and analysis procedures, and to arrive at similar results (Gibbert et al. 2008). Reliability is mostly connected to data collection and processing, which should be both transparent and repeatable. In case studies, the use of case study protocol and the development of a case study database supports reliability by making the procedures taken during the research explicit and all the collected data more easily accessible. The case study protocol is a document that describes the rules and procedures to be followed during a case study, guiding a researcher in conducting a single case study. It is essential for multiple case studies to ensure the data is collected in a similar way in all the individual cases. It also increases the reliability of a study to a significant degree. A case study protocol should give an overview of the project, describe the field procedures and case study questions and provide a guide for producing case study reports or case descriptions of individual cases. A case study database refers to the organisation of the collected data, which is unfortunately often a weak link in qualitative studies. The database should keep the collected data organised, categorised, complete and easily available for later access (Yin 2008). Reliability is especially important because of its relationship with the other validity measures. As Gibbert et al. (2008) argue, researchers who address reliability, and plan against random errors, are more likely to take care of other aspects of validity, such as dealing with systematic errors. Validity measures are not independent of each other either. A clear theoretical and causal logic (internal validity) and links between theoretical concepts and empirical observations (construct validity) are requirements for external validity (Gibbert et al. 2008).

In this study, special attention has been paid to transparency and replicability, the two main ingredients of reliability. Procedures are documented in detail, and original data and intermediate results from the qualitative analysis have been made available, forming a case study database. Statistical analyses are fully repeatable using the (anonymised) datasets and R scripts. Configurations and data are available for the simulation model. Accordingly, this study is reliable, that is, there should be no random errors.

Construct validity means that a study investigates what it claims to investigate. A systematic literature review was used to formulate the theoretical framework for the study. An unobservable abstract concept of collective intelligence was defined in terms of observable concepts, such as rules, tasks, agents, and feedback. Multiple sources of evidence are used including websites, diary notes, content created by users, and statistics. The chain of evidence from raw data to final findings is intact. Data is triangulated from websites, content created by users, news articles, experiences of a single user in the form of diary notes, and statistics. As a result, the constructs of this study are valid and are likely to represent the entities they claim to represent.

Internal validity is about explaining how one thing leads to another. A clear framework is used based on a systematic literature review. Observed patterns are matched to several theoretical frameworks, resulting in theory triangulation. Complementary results from different theoretical lenses corroborate the findings, making the study internally valid.

Fulfilling the above three quality measures is a prerequisite for external validity, which deals with the generalisability of results. Three cases were selected to replicate the results. The cases were compared under different theoretical frameworks. An analogical case was described based on the commonalities between the cases. The results of this study should be valid in other analogical cases such as crowdsourcing platforms where a crowd submits solution suggestions in response to a challenge, evaluates them, and where the crowdsourcing organisation selects the winners.

6.6 Theoretical and practical implications

The presented work has several theoretical implications. First, the systematic literature review and the theoretical framework based on it clarify the concept of collective intelligence. Researchers interested in collective intelligence can use the theoretical framework for organising the growing body of literature, developing testable hypotheses and positioning the new findings. Secondly, discussion on web forums and other related interactions during the creation and refinement of ideas, concepts and designs was identified as a possible manifestation of collective intelligence on crowdsourcing sites.

Thirdly, a form of the wisdom of crowd effect was observed in crowd's evaluations, which is another possible manifestation of collective intelligence on crowdsourcing sites. The aggregated crowd's evaluations were correlated with the expert decisions but in practical settings the wisdom of crowds does not appear to be accurate enough to be relied upon alone in decision making. Finally, feedback loops alone can reduce the accuracy of the crowd, especially if a simple vote count is used to aggregate the evaluations. Although the use of averaging suffers less from feedback loops, the convergence of evaluations may still be slowed. In addition, although two possible manifestations of collective intelligence on crowdsourcing sites were identified, it still appears that there are marked differences to the swarm intelligence of social insects. Individual human beings appear to be the locus of the cognitive work. Emerging intelligence and creativity are mostly functions of individuals. In contrast, the swarm intelligence of insect colonies is more of a group effort, in the sense that insects in a swarm function more like a parts of a hive-mind, a single brain made of insects. As a result, the swarm intelligence of social insects appears to be a scaled-up version of a brain, resulting in cognitive performance that is more than the sum of its parts. Collective intelligence on crowdsourcing sites, however, more closely resembles the combined output of several individual brains, the cognitive performance being merely the sum of the parts. Table 6.3 compares the findings to other examples of group level cognition.

Table 6.3: Examples where collective cognitive capabilities vastly exceed individual capabilities. Findings from this study have been added to the table (the last two rows).

Example	Individual level	Collective level
Nest-site selection of honey bees	Individuals decide whether they like a candidate site	Swarm selects the best available nest site in the environment
Foraging of social insects	Individuals search for food sources, collect food, and advertise food sources to others	Colony optimises foraging among different food sources
Brain	Individual neurons integrate and send signals	Consciousness
Discussion on crowdsourcing sites	Individuals read and write comments	Collective intelligence factor (Woolley et al. 2010; Barlow and Dennis 2014)? Typical human group performance (e.g., Kerr and Tindale 2004)
Evaluation of content on crowdsourcing sites	Individuals evaluate content	Aggregated evaluations are correlated with expert decisions, but are not accurate enough to be relied upon alone in decision making.

The practical implications of this study have an impact on the design of crowdsourcing sites. Additional decision-making mechanisms are needed when using a crowd to evaluate content due to the inaccuracy of crowds in practical settings. Where the accuracy of evaluations is an important factor for the application, feedback loops should be avoided, especially when vote counts are used for aggregating evaluations. Use of averaging reduces the influence of feedback loops to some extent. Uses of crowd’s evaluations other than as support to decision making may be an important factor to consider. Anecdotal evidence suggests that crowd’s evaluations could also be useful for providing feedback to users and for encouraging users to spend time on the website.

7 Conclusions

This study has investigated how crowdsourcing works with a particular look at the extent to which collective intelligence and related phenomena play a part in its functioning. A systematic literature review was conducted and used to develop a theoretical framework of collective intelligence. Three cases of crowdsourcing innovations and product development were investigated from multiple viewpoints in an effort to find phenomena that could be called collective intelligence. On the surface, the innovation processes on all three sites have similar phases and fit the general descriptions of innovation processes found in the literature: a problem is defined more or less explicitly, ideas for possible solutions are generated and evaluated, some of them are selected to be developed further and they finally implemented for use or put up for sale. In other words, the crowdsourcing organisation publishes a challenge and a crowd submits suggestions for solutions. The crowd then evaluates the suggestions, after which the organisation selects winning solutions, possibly using the crowd's evaluations as a guide. The organisation then takes care of implementation or manufacturing of the products or, in case of OpenIDEO, the resulting concepts are available for anyone to implement through a creative commons (Creative Commons 2015) license. Differences between the crowdsourcing sites emerge when looking at the details.

First, using the framework of the collective intelligence genome to analyse the cases revealed three different patterns in the relationships and roles of the crowd and the organisation. OpenIDEO facilitates the creative process of crowd members by giving them tasks and suggestions that help them come up with good concepts. The result is an alternating pattern of activities between the crowd and the organisation. At Threadless, the crowd works mostly independently at the beginning when the crowd members create product designs. Threadless staff take over the process at the very end and only select which designs to print. As the crowd members have already prepared the files to be printed, this is a relatively routine task. The result is a crowd-heavy pattern. At Quirky, the crowd members only come up with the initial idea. After that, members of the Quirky staff carry out most development work. The result is an organisation-heavy pattern, where the crowd participates in small tasks every now and then and the in-house product development team contributes most of the effort.

Secondly, there is a difference in the emphasis on process phases, as revealed by the comparison of inventor's search processes for different sites in Chapter 5 (5.2). OpenIDEO is heavily focused on the early phases of the process and explicitly facilitates the creative process of coming up with ideas for solutions. First, tasks given to users involve investigating the problem or issue at hand in detail, and looking for solutions that have worked elsewhere. Both are known to be valuable approaches to creating good solutions: it is important to understand exactly what is the problem to be solved (Ulrich 2011; IDEO 2009; Christensen et al. 2007). Copying and recombining elements of existing solutions to create new solutions has been accepted as an important and good practice (Murray 2010; Schilling and Green 2011; Kaplan and Vacili 2014). Quirky essentially starts where OpenIDEO stops: participants submit solution ideas and describe the problem they are supposed to solve in the submission form. In principle, these could be winning concepts from OpenIDEO that are ready to be implemented. After selecting concepts, Quirky still has to put lots of effort into developing the concepts into actual products. Threadless becomes involved even later in the development process. Users come up with the ideas, develop them first as concepts (sketches) and may collaborate with other users. Only the finished product design is submitted to Threadless. It is reasonable to assume that each approach requires different skills and capabilities from the crowdsourcing organisation. At OpenIDEO, the focus is on the early phases of the process. Successfully running the crowdsourcing site can be expected to require knowledge on creativity, design, conceptualisation and collaboration. Quirky, on the other hand, focuses on the development phase, which requires in-house design and engineering skills. Not surprisingly, Quirky has dozens of designers and engineers among its staff members. Focus on the implementation phase, as is the case at Threadless, requires manufacturing capabilities and a crowd capable of producing finished designs.

In the rest of the analyses, the cases share similar features. If crowdsourcing innovations are considered as a search problem, the crowdsourcing sites aggregate individual solutions in all cases. Although the phases of the inventor's search process can be identified on the sites, especially at OpenIDEO, individual users still need to act as inventors. In all cases, the enabling properties of collective intelligence systems are easily observed. People are able to adapt, which alone is enough to make the systems as a whole adaptive by definition. All three sites, being interactive websites, allow interaction between the participants. They have both explicit rules and instructions and implicit cultural rules on what is considered acceptable behaviour for the participants. Of the defining properties, randomness, redundancy and robustness are clearly present in all the sites. The behaviour of participants is not deterministic and content is often presented in random order. Multiple participants create redundancy and thus robustness against some participants leaving the site. Identifying collective intelligence, as defined in the framework of collective intelligence systems (Schut 2010), comes down to finding distinctions between the local and global level behaviours of the systems (agents versus the system as a whole). Two possible candidates were identified.

All websites feature discussion forums and there are forum-like discussion features in connection to submitted content. In all cases, participants can comment on submitted concepts, ideas or designs. In addition, the organisations host meetings where they have discussions and make decisions. A consensus emerging from discussions between humans, whether real-life or virtual, could be an example of collective intelligence. More interesting for the current discussion is the emergence of order in the submitted content. Although no single user can view all the content, the crowd is still able to sort the content according to its perceived quality. This is similar to the nest-site selection of honeybees and ant colonies. To investigate the accuracy of evaluations, data from OpenIDEO and Threadless were analysed statistically. Both cases revealed a similar pattern: the crowd is able to predict but cannot categorise. That is, although content with higher scores from the crowd is more likely to be selected by the organisation than content with low scores, the content cannot be accepted or rejected based on the crowd evaluations alone. Often submissions with low scores are selected, and submissions with high scores are not. The crowd cannot be relied upon to select content for further development. A crowd could still be used to filter out the worst content, as Threadless has actually done by dropping low-scoring designs early. Investigating the possible pathways of information on the case sites opens up the possibility for feedback loops on all sites. Content in evaluation can be sorted directly (OpenIDEO and Threadless) or indirectly (Quirky) by current popularity. Popular content could therefore get more attention from the crowd, making it even more popular regardless of the actual quality. Feedback loops may reduce the independence of evaluations and otherwise interfere with the evaluation process. In order to find out what role the feedback loop plays in the functioning of the crowdsourcing systems, a simple agent-based simulation model was created. The simulation was used to compare the effects of different strengths of feedback loops and content evaluation methods on the accuracy of crowd evaluations. Feedback loops alone can delay the convergence of crowd evaluations or even prevent it entirely, depending on the method used to aggregate the crowd's opinion.

7.1 Answers to research questions

The first research question of this thesis was what collective intelligence is. A thorough literature review was conducted, which led to the development of a theoretical framework. Collective intelligence in humans consists of three levels of abstraction. At the micro level, collective intelligence is a combination of psychological, cognitive and behavioural elements that give humans their capabilities for group activities. At the macro level, collective intelligence is best described as a statistical phenomenon, often referred to as the wisdom of crowds effect. Between the micro and the macro level resides the level of emergence, which deals with the question of how behaviour on the macro level emerges from the interactions of individuals at the micro level.

The second question asked how collective intelligence is manifested in websites using crowdsourcing innovation. Three crowdsourcing sites were investigated in detail using mainly participant ethnography and qualitative data analysis methods. Several analytical frameworks were used to take into account different viewpoints on the phenomenon and corroborate findings through triangulation. Two possible manifestations of collective intelligence on crowdsourcing sites were identified: virtual discussions hosted on the websites and the wisdom of crowds during the evaluation of ideas.

The third research question was how important collective intelligence is for the functioning of the crowdsourcing sites. Both manifestations of collective intelligence have some caveats. Sharing and refining knowledge through discussion is typical human behaviour and as such it is not very specific to crowdsourcing sites. Statistical analyses showed that the wisdom of crowds was not accurate enough to be relied upon in decision making. Collective intelligence does not appear to be the major deciding factor for the success or failure of the investigated crowdsourcing sites.

7.2 Contribution

This thesis contributes to the literature on collective intelligence, as defined in Chapter 2 (3.3), by clarifying the concept of collective intelligence and by exploring its role in crowdsourcing in the following ways:

1. Conceptualisation and theoretical framework for collective intelligence based on a systematic literature review.
2. Demonstration on how to use the framework to study collective intelligence.
3. Identification of two possible manifestations of collective intelligence in crowdsourcing sites.
4. Assessment of crowd accuracy in practical settings, and comparison of two typical aggregation methods in the presence of a typical feedback loop.
5. Practical suggestions on how to improve the situation: the crowd should not be allowed to make the final decisions; averaging is preferable to counting votes; feedback loops should be avoided if accuracy is important; and attention should be paid to other uses of crowd evaluations apart from their support for decision making.

Contrary to recent enthusiasm for and speculation on the importance of collective intelligence for crowdsourcing applications, the investigation of three crowdsourcing sites suggests collective intelligence plays only a minor role in explaining the rise and success of crowdsourcing. Reliance on the wisdom of crowds in evaluating content may be novel to crowdsourcing sites, but none of the case examples relied on it solely in their decision making. In addition, the combination of feedback loops and use of the vote counting to aggregate evaluations leaves room for improvement in the accuracy of the crowd.

7.3 Limitations and further research

Several important limitations need to be taken into account when assessing the contributions of this thesis. The initial sample of literature used to develop the theoretical framework was obtained from a single database with only two keyword searches. The scope was limited to papers discussing the collective intelligence of humans. This work could be expanded by reviewing literature from other sources and by including non-human examples. The possibility of mistakes made by the researcher cannot be ruled out. Despite attempting scientific rigour, important sources may have been missed during the cursory analysis of the initial sample. In addition, the identification of the themes and their categorisation is subjective.

Only three similar cases of crowdsourcing innovations were studied, and the results should not be generalised too broadly. Typically for qualitative case studies, the generalisability of results is analogical, that is, similar results can be expected in similar cases. In any situation where a crowd submits, discusses and refines content on a website, evaluates it using either voting or averaging in the presence of feedback loops, and where professional employees select content for further development, the results are likely to hold true. However, for crowdsourcing applications using different approaches to innovation, the situation with regard to collective intelligence may be completely different and the results of this study may not apply.

In ethnographic studies, the researcher acts as an instrument, which introduces possibilities for bias in data collection. The collected data may be more anecdotal than representative. Had the data been collected in a statistically representative way, each page on the websites should have had equal opportunity to be included in the sample. This was not the case. The collected data are rather a sample of what a user might encounter while participating on these sites. As the purpose of the study was to gain understanding of how these complex systems work from the inside, focusing data collection in the areas where users probably spend most of their time on the site is justified. As with ethnographic data collection, the qualitative data analysis may suffer from researcher bias and misunderstanding. This issue was mitigated by maintaining the chain of evidence, using multiple analysis approaches to corroborate findings, and by attempting to keep the analyses transparent. There may be room for interpretation in fitting the details of case descriptions to various analytical frameworks, but the broad conclusions should still hold. Finally, it was unfortunately not possible to check the results of the study with other insiders. The use of other forms of triangulation, such as using both qualitative and quantitative data, and relying on qualitative, statistical and simulation methods in analysis were used to mitigate the risks of misinterpretation.

The findings of this thesis identify three major avenues for further research. First, the theoretical framework of collective intelligence can be used to generate testable hypotheses. This line of research could further clarify the concept and advance the theoretical understanding of collective intelligence. Secondly, new crowdsourcing applications could be developed that aim to replicate the principles of swarm intelligence in the human context more closely. In such an application, participants would perform relatively simple tasks that would contribute to the common innovation goal behind the scenes. This could be achieved by a more fine-grained modularisation of the innovation process, by hiding some of the unnecessary information from users, and by giving more guidance on what the user should do next. Game-like elements could also be added to motivate and direct the users within the application. Thirdly, accuracy could perhaps be improved by using multi-attribute scales, which have been suggested to outperform simple voting, 5-star ratings and prediction markets (Riedl et al. 2010; Blohm et al. 2011). This suggestion is corroborated by extensive evidence on the utility of improper linear models. Improper linear models are linear models where the weights of the predictor variables are obtained in some non-optimal methods, such as intuition, random selection, or by setting them all equal (Dawes 1979). While people are good at selecting useful variables for prediction, they are bad at integrating information from diverse sources. Improper linear models can help with the integration. A classic example of such cases is the Apgar score, which is used to evaluate the condition of new-born babies. To evaluate the condition of an infant, five easily identified characteristics (heart rate, respiration, muscle tone, reflex irritability and skin colour) are assessed and given a score from zero to two. The scores are summed up and the decision on the need for emergency care is made using the total score from 0 to 10 as a guideline. Half a century after its invention, the method is still relevant for predicting the survival of new-born infants (Casey et al. 2001). Hundreds of studies have demonstrated that in most situations improper linear models outperform humans in prediction accuracy (Kahneman 2011).

7.4 Concluding remarks

A doctoral dissertation has been likened to obtaining a driving licence for research. Just as a driver needs to understand different roads and weather conditions to safely steer a car, a researcher needs to know a variety of research methods to effectively address different research questions. For that reason, I have not shied away from learning new tools and approaches during this research project when an opportunity for applying them has presented itself. The methodology should be solid. Care was taken to follow the best practices of case study research, including creating a case study database and maintaining the chain of evidence. In statistical analyses a gold standard of reliability was achieved by providing the data and source code necessary to make the analyses fully repeatable. The current work makes a clear contribution to our knowledge on the role of collective intelligence in crowdsourcing. An interesting, hyped, and somewhat fuzzy concept was clarified and its significance examined within a small sub-group of crowdsourcing applications. As a learning journey the thesis project has been unsurpassable.

References

Abramovitz, M. (1956) Resource and output trends in United States since 1870. *American Economic Review.*

Adamczyk, S., Bullinger, A. and Möslein, K. (2012) Innovation contests: A review, classification and outlook. *Creativity and Innovation Management*, 21, 4.

Afuah, A. and Tucci, C. (2012) Crowdsourcing as a solution to distant search. *Academy of Management Review*, 37, 355-375.

Agar, M. (2004) We have met the other and we're all nonlinear: Ethnography as a nonlinear dynamic system. *Complexity*, 10, 16-24.

Aitamurto, T. (2011) The impact of crowdfunding on journalism: Case study of Spot.us, a platform for community-funded reporting. *Journalism Practice*, 5, 4.

Aitamurto, T., Leiponen, L. and Tee, R. (2011) The promise of idea crowdsourcing - Benefits, contexts, limitations. *Nokia White Paper*.

Aitamurto, T., and Landemore, H. (2013) Democratic participation and deliberation in crowdsourced legislative processes: The case of the law on off-road traffic in Finland. *Communities and Technologies: Workshop on Large-Scale Idea Management and Deliberation Systems*. 2013. ACM.

Alexy, O., Criscuolo, P., and Salter, A. 2011. No Soliciting: Managing Unsolicited Ideas for R&D. *White Paper*, September 2011.

Allison, G. and Zelikov, P. (1999) *Essence of Decision: Explaining the Cuban Missile Crisis (2nd Edition)*. Pearson. 416 pages.

Alonso, O. and Mizzarro, S. (2012) Using crowdsourcing for TREC relevance assessment. I*nformation Processing and Management*, 48.

Amabile, T., Hill, K., Hennessy, B. and Tighe, E. (1994) The work preference inventory: Assessing intrinsic and extrinsic motivational orientations. *Journal of Personality and Social Psychology*, 66, 5, 950–967.

Amazon Mechanical Turk (2015). URL: https://www.mturk.com/mturk/welcome. Accessed 25 Jan 2015.

AnyLogic (2015) Multimethod Simulation Software. URL: http://www.anylogic.com/

Ariely, D., Gneezy, U., Loewenstein, G., Mazar, N. (2011) Large stages and big mistakes. *Federal Reserve Bank of Boston Working Paper*, 05-11. Available at https://www.bostonfed.org/economic/wp/wp2005/wp0511.pdf

Babauta, L. (2009) *The power of less: The fine art of limiting yourself to the essential in business and life*. Hachette Books, 192 pages.

Barlow, J. and Dennis, A. (2014) Not as smart as we think: A study of collective intelligence in virtual groups. *Collective Intelligence 2014*, June 10-12, 2014, Massachusetts Institute of Technology.

Baumeister, R. and Leary, M. (1995) The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 3.

Bergvall-Kåreborn, B. and Howcroft D. (2013) The Apple business model: Crowdsourcing mobile applications. *Accounting Forum*, 37, 280-289.

Blank, S., & Dorf, B. (2012). *The Startup Owner's Manual*. K&S Ranch Press. ISBN-10: 0984999302.

Blohm, I., Riedl, C., Leimeister, J. and Krcmar, H. (2011) Idea evaluation mechanisms for collective intelligence in open innovation communities: Do traders outperform raters? *Thirty Second International Conference on Information Systems*, Shanghai 2011.

Blomqvist, K. (1997) The Many Faces of Trust. *Scandinavian Journal of Management*, 13, 271-286.

Bonabeau, E. (1999) *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, 305 pages.

*Bonabeau, E. (2009) Decisions 2.0: The Power of Collective Intelligence. *MIT Sloan Management Review*, 50, 2: 45-52.

*Bonabeau, E. and Meyer, C. (2001) Swarm Intelligence: A Whole New Way to Think About Business. *Harvard Business Review*, 79, 5: 106-114.

Borshchev, A. (2013) *The Big Book of Simulation Modeling - Multimethod Modeling with AnyLogic 6.* AnyLogic North America. 612 pages.

*Bosse, T., Jonker, C. M., Schut, M. C. and Treur, J. (2006) Collective Representational Content for Shared Extended Mind. *Cognitive Systems Research*, 7, 151-174.

*Bothos, E., Apostolou, D. and Mentzas, G. (2009) Collective Intelligence for Idea Management with Internet-based Information Aggregation Markets. *Internet Research*, 19, 26-41.

Boudreau, K. and Lakhani, K. (2009) How to manage outside innovation. *MIT Sloan Management Review*, 50, 4.

Boudreau, K., Lacetera, N. and Lakhani, K. (2011) Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis. *Management Science*, 57, 5.

Brabham, D. (2008a) Moving the crowd at iStockphoto: The composition of the crowd and motivations for participation in a crowdsourcing application. *First Monday*, 13.

Brabham, D. (2008b) Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence,* 14, 75-90.

*Brabham, D. C. (2009) 'Crowdsourcing the Public Participation Process for Planning Projects,' Planning Theory, 8, 242-262.

*Brabham, D. C. (2010) Moving the Crowd at Threadless. *Information, Communication & Society*, 13, 1122-1145.

Bryman, A. and Bell, E. (2007) Business Research Methods. Oxford University Press.

Bullinger, A.C. and Möslein, K.M. (2010) Innovation contests – Where are we? *Proceedings of the 16th Americas Conference on Information Systems*.

*Cachia, R., Compañó, R. and Da Costa, O. (2007) Grasping the Potential of Online Social Networks for Foresight. *Technological Forecasting & Social Change*, 74, 1179-1203.

Camazine, S., Deneubourg, J-L., Franks, N. R., Sneyd, J., Theraulaz, G and Bonabeau, E. (2001) *Self-organization in Biological Systems*, Princeton University Press, 560 pages.

Cantisani, A. (2006) Technological innovation processes revisited. *Technovation*, 26, 1294-1301.

Casey, B., McIntire, D. and Leveno, K. (2001) The continuing value of the APGAR score for the assessment of newborn infants. *The New England Journal of Medicine*, 344, 7.

Castellion, G. and Markham, S. (2012) Perspective: New product failure rates: Influence of Argumentum ad Populum and self-interest. *Journal of Product Innovation Management*, 30, 5)

Chakravartty, A. (2014) Scientific Realism. *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL: http://plato.stanford.edu/archives/spr2014/entries/scientific-realism/.

*Chen, C. (2007) Holistic Sense-making: Conflicting Opinions, Creative Ideas, and Collective Intelligence. *Library Hi Tech*, 25, 311-327.

Cherniss, C. (2010) Emotional Intelligence: Toward Clarification of a Concept. *Industrial and Organizational Psychology-Perspectives on Science and Practice*, 3, 110-126.

Chesbrough, H. (2003) *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Review Press. 272 pages.

Christensen, C., Anthony, S., Berstell, G. and Nitterhouse, D. (2007) Finding the right job for your product. *MIT Sloan Management Review*, 49, 3.

*Coe, A., Paquet, G. and Roy, J. (2001) E-governance and Smart Communities: A Social Learning Challenge. *Social Science Computer Review*, 19, 80-93.

Coe, R. (2002) It’s effect size, stupid - What effect size is and why it is important. *British Educational Research Association annual conference*, Exeter, 1214 September, 20

Conradt, L. and List, C. (2009) Group decisions in humans and animals: a survey. *Philosophical Transactions of the Royal Society B*, 364.

Conradt, L. and Roper, T. (2005) Consensus decision making in animals. *Trends in Ecology and Evolution.* 20: 8.

Cooper, R. (1990) Stage-Gate Systems: A New Tool for Managing New Products. *Business Horizons*, 33, 44-54.

Cooper, R., Edgett, S. ad Kleinschmidt, E. (2002) Optimizing the State-Gate Process: What Best-practice Companies Do – I, *Research Technology Management*, 45, 21-27.

Creative Commons (2015) Creative Commons website. URL: https://creativecommons.org/ Accessed 16 April 2015.

Crowdsourcing.org (2015) Crowdsourcing.org The Industry Website. URL: http://www.crowdsourcing.org/. Accessed 25 Jan 2015.

Cummings, S., Daellenbach, U., Davenport, S. and Campbell, C. (2012) "Problem-sourcing”: a re-framing of open innovation for R&D organizations. *Management Research Review*, 36, 10.

Damper, R. I. (2000) Emergence and Levels of Abstraction. *International Journal of Systems Science*, 31, 811-818.

Dawes, R. (1979) The robust beauty of improper linear models in decision making. *American Psychologist*, 34, 7.

Deci, E., Koestner, R. and Ryan, R. (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125, 6.

Dedoose (2013) Dedoose. URL: http://www.dedoose.com/. Accessed 18.2.2013.

Desouza, K., Dombrowski, C., Awazu, Y., Baloh, P., Papagari, J., Jha S. and Kim, J. (2009) Crafting organizational innovation processes. *Innovation: Management, Policy & Practice*, 11, 6-33.

De Wolf, T. and Holvoet, T. (2005) Emergence versus self-organization: different concepts but promising when combined. in: S. Brueckner, G.D.M. Serugendo, A. Karageorgos, R. Nagpal (Eds.), *Proceedings of the Workshop on Engineering Self Organising Applications, vol. 3464 of Lecture Notes in Computer Science*, Springer, 2005, pp. 1–15

Di Gangi, P. and Wasko, M. (2009) Steal my idea! Organizational adoption of user innovations from a user innovation community: A case study of Dell IdeaStorm, *Decision Support Systems*, 48, 303-312.

Doan, A., Ramakhrisnan, R. and Halevy, A. (2011) Crowdsourcing Systems on the World-wide Web. *Communications of the ACM*, 54, 86-96.

Dorigo, M., Bonabeau, E. and Theraulaz, G. (2004) Ant algorithms and stigmergy. *Future Generation Computer Systems*, 16.

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

Eisenhardt, K. (1989) Building theories from case study research. T*he Academy of Management Review*, 14, 532-550.

Eisenhardt, K. and Graebner, M. (2007) Theory building from cases: opportunities and challenges. *Academy of Management Review*, 50, 1.

European Commission (2010) Europe 2020: A strategy for smart, sustainable and inclusive growth. *Communication from the Commission.* URL: http://www.europeanpaymentscouncil.eu/index.cfm/knowledge-bank/other-sepa-information/european-commission-communication-europe-2020-a-strategy-for-smart-sustainable-and-inclusive-growth-march-2010-/

Evernote (2013a) Evernote. URL: http://evernote.com/evernote/. Accessed 18.2.2013.

Evernote (2013b) Evernote Web Clipper. URL: http://evernote.com/webclipper/. Accessed 18.2.2013

*Fathianathan, M., Panchal, J. H. and Nee, A. Y. C. (2009) A Platform for Facilitating Mass Collaborative Product Realization. *CIRP Annals – Manufacturing Technology*, 58, 127-130.

Feynman, R., Leighton, B and Sands, M. (1964) *The Feynman Lectures in Physics*.

Fenn, D. (2011) The man behind Quirky. *Inc Magazine*, April 2011. URL: http://www.inc.com/magazine/20110401/the-man-behind-quirky.html. Accessed: 15 May 2013.

Field, A., Miles, J. and Field, Z. (2012) *Discovering Statistics Using R*. Sage Publications, 957 pages. ISBN: 978-1446200469

Finnerty, A., Kucherbaev, P., Tranquillini, S. and Convertino, G. (2013) Keep it simple: Reward and task design in crowdsourcing. *CHItaly '13*, September 16 - 20 2013, Trento, Italy

Fleming, L. and Sorenson, O. (2001). Technology as a complex adaptive system: Evidence from patent data. *Research policy*, 30, 1019-1039.

*Franck, G. (2002) The Scientific Economy of Attention: A Novel Approach to Collective Rationality of Science. *Scientometrics*, 55, 3-26.

Frey, C. and Osborne, M. (2013) The future of employment: How susceptible are jobs to computerisation? *Working paper*.

Gallagher, S. (2013) The socially extended mind. *Cognitive Systems Researc*h.

*Gan, Y. and Zhu, Z. (2007) A Learning Framework for Knowledge Building and Collective Wisdom Advancement in Virtual Learning Communities. *Educational Technology & Society*, 10, 206-226.

Gabelloni, D., Ghezzi, A. and Martini, A. (2013) Mapping the research on crowdsourcing: an interpretive review and research agenda. *CINet 2013*.

Geary, D. (2005) *Origin of Mind: Evolution of brain, cognition and general intelligence*, American Psychological Association, 459 pages.

Gerstenfeld, P., Grant, D. and Chiang, C-P. (2003) Hate online: A content analysis of extremist internet sites. *Analyses of Social Issues and Public Policy,* 3, 1.

Gibbert, M., Ruigrok, W. and Wicki, B. (2008) What passes as a rigorous case study? *Strategic Management Journal*, 29, 1465-1474.

Girotra, K., Terwiesch, C. and Ulrich, K. (2010) Idea generation and the quality of the best idea. *Management Science*, 56, 4.

Godfrey-Smith, P. (2003) *Theory and reality: An introduction to philosophy of science*. University of Chicago Press. 272 pages.

Goldstone, R. and Gureckis, T. (2009) Collective behavior. *Topics in Cognitive Science*, 1.

Gordon, D., Holmes, S. and Nacu, S. (2008) Short-term regulation of foraging in harvester ants. *Behavioral Ecology*, 19, 217-222.

*Gregg, D. (2009) Developing a Collective Intelligence Application for Special Education. *Decision Support Systems*, 47, 455-465.

*Gregg, D. (2010) Designing for Collective Intelligence. *Communications of the ACM*, 53, 134-138.

Griffith, E. (2013) With $79 million Series D round, Quirky is New York's most underrated (yet well-funded) startup. *Pando Daily*. URL: http://pando.com/2013/11/13/with-79-million-series-d-round-quirky-is-new-yorks-most-underrated-yet-well-funded-startup/. Accessed: 12 April 2015.

*Gruber, T. (2008) Collective Knowledge Systems: Where the Social Web Meets the Semantic Web. *Journal of Web Semantics*, 6, 4-13.

Güney, S. (2010) New significance for an old method: CAS theory and ethnography, *Communication Methods and Measures*, 4, 273-289.

Gyatso, T (2005) Our faith in science. *The New York Times*, 12 November 2005. URL: http://www.nytimes.com/2005/11/12/opinion/12dalai.html?pagewanted=all&_r=0. Accessed 17 April 2015.

Fenn, D. (2012) Quirky gets backing from Andreesse Horowitz and Kleiner Perkins. *The New York Times*, 7 September 2012. URL: http://dealbook.nytimes.com/2012/09/07/quirky-gets-backing-from-andreessen-horowitz-and-kleiner-perkins/. Accessed 15 May 2013.

Heylighen, F. (1999) Collective Intelligence and its Implementation on the Web: Algorithms to Develop a Collective Mental Map. *Computational & Mathematical Organization Theory*, 5, 253-280.

Heylighen, F. (2013) Self-organization in communicating groups: The emergence of coordination, shared references and collective intelligence. *Understanding Complex Systems*, 117-149.

Heylighen, F. (2011) Conceptions of a Global Brain: an historical review. In *Evolution: Cosmic, Biological, and Social*, eds. Grinin, L. E., Carneiro, R. L., Korotayev A. V., Spier F. (pp. 274 – 289). Uchitel Publishing. Retrieved from http://pcp.vub.ac.be/papers/GB-conceptionsRodrigue.pdf

Hippel, E. (1986) Lead users: A source of novel product concepts. *Management Science*, 32, 7.

Hippel, E. (2005) *Democratizing Innovation*. The MIT Press. Available at http://web.mit.edu/evhippel/www/democ1.htm

Hollan, J., Hutchins, E. and Kirsh, D. (2000) Distributed cognition: Towards a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, Vol. 7, No. 2, June 2000, Pages 174–196.

Hong, L. and Page, S. (2004) Groups of Diverse Problem-solvers Can Outperform Groups of High-ability Problem-solvers. *PNAS*, 101, 16385-16389.

Howe, J. (2006) The rise of crowdsourcing. *Wired Magazine*. June 2006.

Howe, J. (2008) *Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business*. Random House Books, 312 pages.

Hoyer, W., Chandy, R., Dorotic, M., Krafft, M. and Singh, S. (2010) Consumer cocreation in new product development. *Journal of Service Research*, 13, 3.

Hudson, C. (2015) 100 30 day challenge ideas to turn your life around. *ThoughtBrick*, 5 February 2015. URL: http://thoughtbrick.com/lifestyle/hundred-30-day-challenge-ideas/. Accessed 17 April 2015.

Huston, L. and Sakkab, N. (2006) Connect and develop: Inside Procter & Gamble's new model for innovation. *Harvard Business Review*.

Hutchins, E. (1995) How a cockpit remembers its speed. *Cognitive Science*, 19, 265-288.

Hutchins, E. (1996) *Cognition in the Wild*. Bradford Books. 408 pages.

Hutchins, E. and Klausen, T. (1995) Distributed cognition in an airline cockpit. In Y. Engeström and D. Middleton (Eds.) *Cognition and communication at work*. New York: Cambridge University Press. pp. 15-34.

IDEO (2009) *Human Centered Design Toolkit 2nd Edition*, IDEO. http://www.ideo.com/work/item/human-centered-design-toolkit/ Accessed 5 August 2010.

InnoCentive (2013) About Us. URL: http://www.innocentive.com/about-innocentive. Accessed 15 Jan 2013.

Ioannidis, J. (2005) Why most published research findings are false. *PLOS Medicine*, 2, 8.

IPCC (2014) Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

James, G., Witten, D., Hastie, T. and Tibshirani, T. (2013) *An Introduction to Statistical Learning with Applications in R*. Springer.

Jeppesen, L. and Lakhani, K. (2010) Marginality and problem-solving effectiveness in broadcast search, *Organization Science*, 21, 1016-1033.

Johnson, S. (2010) *Where do good ideas come from: The natural history of innovation*. Riverhead Books. 336 pages.

Johnson, V. (2013) Revised standards for statistical evidence. *PNAS*. 110, 48, 19313-19317.

*Jung, Y., Park, Y., Bae, H. J. and Lee, B. S. (2011) Employing Collective Intelligence for User Driven Service Creation. *IEEE Communications Magazine*, 49, 75-83.

Just, C. (2011) A Review of Literature on General Factor of Personality. *Personality and Individual Differences*, 50, 765-771.

Kahneman, D. (2011) *Thinking, Fast and Slow*. Farrar, Straus and Giroux. 512 pages.

Kaplan, S. and Vakili, K. (2014) The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal.*

Kauffman, S. (1993) *The Origins of Order: Self-organization and selection in evolution*, Oxford University Press, 728 pages.

Kauffman, S., Lobo, J. and Macready, W. (2000). Optimal Search on a Technology Landscape, *Journal of Economic Behavior & Organization*, 43, 141-166.

Kerr, N. L. and Tindale, R. S. (2004) Group Performance and Decision Making, *Annual Review of Psychology*, 55, 623-655.

Khatib, F., DiMalo, F., FoldIt Contenders Group, FoldIt Void Crushers Group, Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., Zabranska, H., Pichova, I., Thompson, J., Popovic, Z., Jaskolski, M. and Baker, D. (2011) Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature Structural & Molecular Biology*, 18, 1175-1177.

Kittur, N., Nickerson, J., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M., and Horton, J. (2013) The future of crowd work. *CSCW '13*, February 23–27, 2013, San Antonio, Texas, USA.2

Knutas, A., Hajikhani, A., Salminen, J., Ikonen, J. and Porras, J. (2015) Cloud-based bibliometric analysis service for systematic mapping studies. *International Conference on Computer Systems and Technologies*, 25-26 June 2015, Dublin.

Kohn, N., Paulus, P. and Choi, Y. (2011) Building on the ideas of others: An examination of the idea combination process. *Journal of Experimental Social Psychology*.

*Komninos, N. (2004) Regional Intelligence: Distributed Localised Information Systems for Innovation and Development. I*nternational Journal of Technology Management*, 28, 483-506.

Kozinets, R. (2002) The Field behind the screen: Using netnography for marketing research in online communities. *Journal of Marketing Research*, 39.

Kozinets, R. (2010) *Netnography: Doing Ethnographic Research Online*. Sage Publigations. 232 pages.

*Krause, J., Ruxton, G. and Krause, S. (2009) Swarm Intelligence in Animals and Humans. *Trends in Ecology and Evolution*, 25, 28-34.

*Krause, S., James, R., Faria, J. J., Ruxton, G. D. and Krause, J. (2011) Swarm Intelligence in Humans: Diversity Can Trump Ability. *Animal Behaviour*, 81, 941-948.

Kumar, V. (2009) A Process for Practicing Design Innovation. *Journal of Business Strategy*. 30, 91-100.

Lampel, J., Jha, P. and Bhalla, A. (2012) Test-driving the future: How design competitions are changing innovation. *Academy of Management Perspectives*, 26, 2.

Landemore, H. (2014) Inclusive constitution-making: The Icelandic experiment. *The Journal of Political Philosophy*.

Lansing, J. S. (2003) Complex adaptive systems. *Annual Review of Anthropology*, 32.

*Lee, J-H. and Chang, M-L. (2010) Stimulating Designers' Creativity Based on a Creative Evolutionary System and Collective Intelligence in Product Design. *International Journal of Industrial Ergonomics*, 40, 295-305.

*Leimeister, J. M. (2010) Collective Intelligence. B*usiness & Information Systems Engineering*, 2, 245-248.

Levinthal, D. 1997. Adaptation on Rugged Landscapes. *Management Science*, 43, 934-950.

Levy, P. (1997) *Collective Intelligence: Mankind's Emerging World in Cyberspace*. Basic Books, 277 pages.

*Levy, P. (2010) From Social Computing to Reflexive Collective Intelligence: The IEML Research Program. *Information Sciences*, 180, 71-94.

*Lorenz, J., Rauhut, H., Schweitzer, F. and Helbing, D. (2011) How Social Influence Can Undermine the Wisdom of Crowd Effect. *PNAS*, 108, 9020-9025.

*Luo, S., Xia, H., Yoshida, T. and Wang, Z. (2009) Toward Collective Intelligence of Online Communities: A Primitive Conceptual Model. *Journal of Systems Science and Systems Engineering*, 18, 2: 203-221.

*Lykourentzou, I., Papadaki, K., Verkados, D. J., Polemi, D. and Loumos V. (2010) CorpWiki: A Self-regulating Wiki to Promote Corporate Collective Intelligence Through Expert Peer Matching. *Information Sciences*, 180, 18-38.

Mackay, C. (1841) *Memoirs of extraordinary popular delusions and madness of crowds*. Richard Bentley, London.

Maggitti, P., Smith, K. and Katila, R. 2013. The complex search process of invention. *Research Policy*, 42, 90-100.

*Malone, T. W., Laubacher, R. and Dellarocas, C. (2010) The Collective Intelligence Genome. *MIT Sloan Management Review*, 51, 3: 21-31.

Malone, T.W., Laubacher, R. and Johns, T. (2011) The age of hyperspecialization. *Harvard Business Review*.

Manning PL, Payne D, Pennicott J, Barrett PM, Ennos RA (2006) Dinosaur killer claws or climbing crampons? *Biol Lett*, 22: 110–112.

Marchi, G., Giaghetti, C. and Gennaro, P. (2011) Extending lead-user theory to online brand communities: The case of the community Ducati. *Technovation*, 31.

Massanari, A. (2012) DIY design: How crowdsourcing sites are challenging traditional graphic design practice. *First Monday*, 17.

McFadzean, E., O'Loughlin, A. and Shaw, E. (2005) Corporate entrepreneurship and innovation part 1: the missing link. *European Journal of Innovation Management*, 8, 350-372.

Merriam-Webster (2015) Agent. URL: http://www.merriam-webster.com/dictionary/agent. Accessed 24 September 2015.

Merrit, M., Stapleton, M. and Huebner, B. (2013) Socially embedded cognition. *Cognitive Systems Research*.

Miles, M. and Huberman, M. (1994) *Qualitative Data Analysis: An Expanded Sourcebook (2nd Edition)*, SAGE Publications, 352 pages.

MIT Center for Collective Intelligence (2015) URL: http://cci.mit.edu/. Accessed 23 March 2015.

Mondada, F., Pettinaro, G., Guignard, A., Kwee, I., Floreano, D., Deneoubourg, J-L., Nolfi, S., Gambardella, L. and Dorigo, M. (2004) Swarm-bot: A new distributed robotic concept. *Autonomous Robots*, 17, 193-221.

Motoyama, M., Levchenko, K., Kanich, C., McCoy, D., Voelker, G. and Savage, S. (2010) Re:CAPTCHAs - Understanding CAPTCHA-solving services in an economic context. *USENIX Security'10 Proceedings of the 19th USENIX conference on Security*.

Mortara, L., Ford, S. and Jaeger, M. (2013) Idea competitions under scrutiny: Acquisition, intelligence or public relations mechanism? *Technological Forecasting and Social Change*, 80.

Muchnik, L., Aral, S. and Taylor, S. (2013) Social influence bias: A randomized experiment. *Science*, 341, 647.

Mullen, B., Johnson, C. and Salas, E. (1991) Productivity loss in brainstorming groups: A meta-analytic integration. *Basic and Applied Social Psychology*, 12, 1.

Murray, D. (2010) *Borrowing Brilliance: The six steps to business innovation by building on the ideas of others*. Gotham. 304 pages.

Murty, P., Paulini, M., Maher, M. (2010) Collective intelligence and design thinking. DTRS'10: Design Thinking Research Symposium, Sydney, Australia.

Myers, D. (2012) *Exploring social psychology (6th ed.)* McGraw-Hill, New York.

Nickerson, R. (1998) Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2, 2.

*Nguyen, N. T. (2008) Inconsistency of Knowledge and Collective Intelligence. *Cybernetics and Systems*, 39, 542-562.

Onarheim, B. and Christensen, B. (2012) Distributed idea screening in stage-gate development processes. *Journal of Engineering Design*, 23, 9, 660-673.

OpenIDEO (2012). About Us. URL: http://www.openideo.com/about-us. Accessed 17 July 2012.

OpenIDEO (2012b) Frequently Asked Questions. URL: https://openideo.com/faq. Accessed 27 July 2012.

OpenIDEO (2015) Our Guiding Principles. URL: https://openideo.com/content/our-guiding-principles. Accessed: 5 June 2015.

Ottino, J. M. (2004) Engineering Complex systems. *Nature*, 427, 339.

Page, S. (2007) T*he Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies*. Princeton University Press, 448 pages.

Page, S. (2012) Model Thinking. Coursera. URL: https://www.coursera.org/course/modelthinking

Partanen, M. (2011) Personal communication.

Passino, K. and Seeley, T. (2005) Modeling and analysis of nest-site selection by honeybee swarms: the speed and accuracy trade-off. *Behavioral Ecology and Sociobiology.* 59, 427-442.

Passino, K., Seeley, T. and Visscher, P. (2008) Swarm Cognition in Honey Bees. *Behavioral Ecology and Sociobiology*, 62, 401-414.

Paulini, M. (2012) Collective intelligence in online innovation communities. *A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy.* The University of Sydney.

Paulini, M., Maher, M. and Murty, P. (2011) The role of collective intelligence in design. In *C. M. Herr, N. Gu, S. Roudavsky, M. A. Schnabel (eds.), Circuit Bending, Breaking and Mending: Proceedings of the 16th International Conference on Computer-Aided Architectural Design Research in Asia CAADRIA 2011*.

PWC (2011) Growth reimagined - Prospects in emerging markets drive CEO confidence. 14th Annual Global CEO Survey. URL: http://www.pwc.com/gx/en/ceo-survey/pdf/pwc_14th_annual_global_ceo_survey_e.pdf. Accessed: 20 March 2015.

PWC (2012) Delivering results - Growth and value in volatile world. 15th Annual Global CEO Survey. URL: http://www.pwc.com/gx/en/ceo-survey/pdf/15th-global-pwc-ceo-survey.pdf. Accessed: 20 March 2015.

Peng, R. (2009) Reproducible research and biostatistics. *Biostatistics*, 10, 3.

*Pentland, A. (2006) Collective Intelligence. *IEEE Computational Intelligence Magazine*, 1, 3: 9-12.

*Pentland, A. (2007) On the Collective Nature of Human Intelligence. *Adaptive Behavior*, 15, 2: 189-198.

Pisano, G. and Verganti, R. (2008) Which kind of collaboration is right for you? *Harvard Business Review*.

Poetz, M. and Schreier, M. (2012) The value of crowdsourcing: Can users really compete with professionals in generating new product ideas? *Journal of Product Innovation Management*, 29, 2.

Pomerlau, D. (2009) Twitter and the Global Brain. Retrieved 1 June 2011 from https://docs.google.com/Doc?docid=0AT9namoRlUKMZGM3eG40ZHNfN2Z3ajJ0eGNw&hl=en&pli=1

Pässilä, A. and Oikarinen, T. (2011) Research-based theatre as a facilitator for organisational learning. In Meusburger, A.P., Berthoin-Antal, A. and Wunder, E. (eds.) *Knowledge in organizations - Learning organizations. Vol. 6 of the series Learning, Knowledge and Space*. Springer.

Quirky (2013) The Quirky website. URL: https://www.quirky.com/.

Quirky (2013a) Inventing for Wink: Instantly connected? Here are some tips… The Quirky blog. URL: https://www.quirky.com/blog/post/2013/04/want-to-invent-for-wink-instantly-connected-here-are-some-tips/ Accessed 19 May 2013.

Quirky (2013b) Weekly Recap 4/13-4/19. The Quirky Blog. URL: https://www.quirky.com/blog/post/2013/04/weekly-recap-413-419/. Accessed 22. April 2015.

Quirky (2013c) About us. URL: https://www.quirky.com/about. Accessed 14 May 2013.

Quirky (2013d) This week's brainstorm review. URL: http://www.quirky.com/blog/post/2013/04/this-weeks-brainstorm-preview-39/. Accessed: 9 May 2013.

Quirky (2013e) How to invent with Quirky. URL: https://www.quirky.com/learn. Accessed: 19 May 2013.

Quirky (2013f) More on voting! Quirky forum discussion. URL: http://community.quirky.com/#filters%5B%5D=all&filters%5B%5D=recent&infinite=false&options=true&page=3&per_page=15. Accessed: 22 May 2013.

Quirky (2013g) No brainstorm, but a spring cleaning feedback Friday! URL: http://www.quirky.com/blog/post/2013/04/no-brainstorm-but-a-spring-cleaning-feedback-friday/. Accessed: 9 May 2013.

Quirky (2013h) Quirky ideas. URL: https://www.quirky.com/ideas/new. Accessed: 3 September 2012.

*Rasmussen, S., Raven, M. J., Keating, G. N. and Bedau, M. A. (2003) Collective Intelligence of the Artificial Life Community on Its Own Successes, Failures, and Future. *Artificial Life*, 9, 207-235.

R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Reid, E. (2013) Crowdsourcing and gamification techniques in Inspire (AQAP online magazine). *ISI 2013,* June 4-7, 2013, Seattle, Washington, USA.

Riedl, C., Blohm, I., Leimeister, J. and Krcmar H. (2010) Rating scales for collective intelligence in innovation communities: Why quick and easy decision making does not get it right. *Thirty First International Conference on Information Systems*, St. Louis 2010.

Rosenberg, N. (2003) Innovation and economic growth. *OECD Conference on Innovation and Growth in Tourism in Lugano*, Switzerland, 18-19 September 2003.

Rothwell, R. (1994) Towards the fifth-generation innovation process. *International Marketing Review*, 11, 1.

Russell, S. and Norvig, P. 1995. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 932 pages.

Salganik, M. and Watts, D. (2008) Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social Psychology Quarterly*, 71:4, 338-355.

Salganik, M. and Watts, D. (2009) Web-based experiments for the study of collective social dynamics in cultural markets. *Topics in Cognitive Science*, 1.

Salganik, M. and Levy, K. (2012) Wiki surveys: Open and quantifiable social data collection. *Technical Report* arXiv:1202.0500, 2012.

SpaceX (2014) X marks the spot: Falcon 9 attempts ocean platform landing. URL: http://www.spacex.com/news/2014/12/16/x-marks-spot-falcon-9-attempts-ocean-platform-landing. Accessed 16 April 2015.

Salminen, J. (2012) Collective intelligence in humans: A literature review, *Collective Intelligence 2012*, Boston, USA. 18.-20.4.2012.

Salminen, J. (2013a) Collective intelligence on a crowdsourcing site. *GBI Working Paper*. Available at https://sites.google.com/site/gbialternative1/working-papers.

Salminen, J. (2013b) Crowdsourcing innovations as a search process. *PIN-C 2013*, Lahti, Finland. 18.-20. June 2013.

Salminen, J. (2014) Wisdom of crowds in practice. *Collective Intelligence 2014*, Boston, USA. 10.-12. June 2014.

Salminen, J. (2015) GitHub account. URL: https://github.com/jsalminen/thesis.

*Scarlat, E. and Maries, I. (2009) Increasing Collective Intelligence within Organizations Based on Trust and Reputation Models. *Economic Computation and Economic Cybernetics Studies and Research*, 43, 2: 61-72.

Schilling, M., and Green, E. 2011. Recombinant Search and Breakthrough Idea Generation: An Analysis of High Impact Papers in the Social Sciences. *Research Policy*, 40, 1321-1331.

Schirr, G. (2012) Flawed tools: The efficacy of group research methods to generate customer ideas. *Journal of Product Innovation Management*, 29, 3.

*Schut, M. C. (2010) On Model Design for Simulation of Collective Intelligence. *Information Sciences*, 180, 132-155.

Scekic, O., Truong, H-L. and Dustdar, S. (2013) Incentives and rewarding in social computing. *Communications of the ACM*, 56, 6.

Seeley, T. D. and Buhrman, S. C. (1999) Group Decision Making in Swarms of Honey Bees. *Behavioral Ecology and Sociobiology*, 45, 19-31.

Seeley, T., Visscher, P., Passino, K. (2006) Group decision making in honey bee swarms. *American Scientist*, 94, 220-229.

Shaw, E., O’Loughlin, A. and McFazdean, E. (2005) Corporate entrepreneurship and innovation part 2: a role- and process-based approach. *European Journal of Innovation Management*, 8, 393-408.

*Shimazu, H. and Koike, S. (2007) KM 2.0: Business Knowledge Sharing in the Web 2.0 age. *NEC Technical Journal*, 2, 2: 50-54.

Sibbet, D. (2010) *Visual Meetings: How Graphics, Sticky Notes and Idea Mapping Can Transform Group Productivity*, Wiley, 288 pages.

Siggelkow, N. (2007) Persuasion with case studies. *Academy of Management Review*, 50, 1.

Simula, H. and Vuori, M. (2012) Benefits and barriers of crowdsourcing in B2B firms: Generating ideas with internal and external crowds. *International Journal of Innovation Management*, 16, 6.

Sousa, T., Silva, A. and Neves, A. (2004) Particle swarm based data mining algorithms for classification tasks. *Parallel Computing*, 30.

Spurlock, M. (2004) *SuperSize Me*. The Con.

Steffen, W., Richardson, K., Rockström, J., Cornell, S., Fetzer, I., Bennett, E., Riggs, R., Carpenter, S., de Vries, W., de Wit, C., Folke, C., Gerten, D., Heinke, J., Mace, G., Persson, L., Ramanathan, V., Reyers, B. and Sörlin, S. (2015) Planetary boundaries: Guiding human development on a changing planet. *Science*, 347, 6223.

Sterman, J. (2000) *Business Dynamics: Systems Thinking and Modeling for a Complex World*. McGraw-Hill. 1008 pages.

Sullivan, E. (2010) A Group effort: More companies are turning to the wisdom of the crowd to find ways to innovate. *Marketing News*, February 28, 22-28.

Surowiecki, J. (2005) *Wisdom of Crowds*. Anchor Books, 306 pages.

Sutton, R. and Hargadon, A. (1996) Brainstorming groups in context: Effectiveness in product design firm. *Administrative Science Quarterly*, 41, 4.

*Szuba, T. (2001) A Formal Definition of the Phenomenon of Collective Intelligence and Its IQ Measure. *Future Generation Computer Systems*, 17, 489-500.

*Szuba, T. (2002) Universal Formal Model of Collective Intelligence and Its IQ Measure. *Lecture Notes in Artificial Intelligence*, 2296, 303-312.

Terwiesch, C. and Xu, Y. (2008) Innovation contests, open innovation, and multi-agent problem-solving. *Management Science*, 54, 1529-1543.

Tidd, J., Bessant, J. and Pavitt, K. (2005) *Managing innovation: Integrating Technological, Market and Organizational Change*. John Wiley & Sons.

Thagard, P. and Stewart, T. (2014) Two theories of consciousness: Semantic pointer competition vs. information integration, *Consciousness and Cognition*, 30, 73-90.

Theraulaz, G. and Bonabeau, E. (1999) A Brief History of Stigmergy, *Artificial Life*, 5, 97-116.

Threadless (2013a) Frequently Answered Questions. Retrieved 7.10.2013. URL: http://support.threadless.com/ics/support/default.asp?deptID=15140&_referrer=http://www.threadless.com/.

Threadless (2013b) Threadless Challenge. Retrieved 1 July 2013. URL: http://www.threadless.com/threadless/

*Trianni, V., Tuci, E., Passino, K. M. and Marshall, J. A. R. (2011) Swarm Cognition: an Interdisciplinary Approach to the Study of Self-organizing Biological Collectives. *Swarm Intelligence*, 5, 3-18.

Tufte, E. (2009) *Visual Display of Quantitative Information, 2nd Edition*. Cheshire Graphics Press. 200 pages.

Turner, J. S. (2011) Termites as Models of Swarm Cognition. *Swarm Intelligence*, 5, 19-43.

Ulrich, K. (2011) *Design: Creation of Artifacts in Society*. University of Pennsylvania, 184 pages.

*Vanderhaeghen, D. and Fettke, P. (2010) Organizational and Technological Options for Business Process Management from the Perspective of Web 2.0: Results of a Design Oriented Research Approach with Particular Consideration of Self-Organization and Collective Intelligence. *Business & Information Systems Engineering*, 2, 15-28.

Veryzer, R. (1998) Discontinuous Innovation and the New Product Development Process. *Journal of Product Innovation Management*, 15, 304-321.

Visscher, K. (2007) Group decision making in nest-site selection among social insects. *Annual Review of Entomology*, 52, 255-275.

Vuculescu, O. and Bergenholtz, C. (2014) How to solve problems with crowds: A computer-based simulation model. *Creativity and Innovation Management*, 23, 2.

Vuurens, J. and de Vries, A. (2012) Obtaining high-quality relevance judgments using crowdsourcing. *IEEE Internet Computing*.

Wainer, H. (1984) How to display data badly. *The American Statistician*, 38, 2.

Walcher, D. (2007) Der Ideenwettbewerb als Methode der aktiven Kundenintegration. Gabler Verlag, Wiesbaden.

Wagner, E. (2011) Why prize? The surprising resurgence of prizes to stimulate innovation. *Research Technology Management*.

Wang, J. (2011) Quirky: The solution to innovator's dilemma. *Entrepreneur*, 26 July 2013. URL: http://www.entrepreneur.com/article/220045. Accessed 15 May 2013.

Web of Science Core Collection (2014) URL: http://webofscience.com. Accessed 9 December 2014.

Wicham, H. (2011) Tidy data. *Journal of Statistical Software*, 59, 10.

Wilkinson, L. and Task Force on Statistical Inference (1999) Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 8.

*Woodley M. A., and Bell, E. (2011) Is Collective Intelligence (mostly) the General Factor of Personality? A Comment on Woolley, Chabris, Pentland, Hashmi and Malone (2010). *Intelligence*, 39, 79-81

*Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N. and Malone, T. (2010), Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*, 330, 686-688.

Woolley, A. W. and Fuchs, E. (2011) Collective Intelligence in the Organisation of the Science. *Organization Science*, 22, 1359-1367.

*Wu, J. and Aberer, K. (2003) Swarm Intelligent Surfing in the Web. *Lecture Notes in Computer Science*, 2722, 431-440.

Yin, R. (2008) *Case Study Research: Design and Methods*, SAGE Publications, 240 pages.

Yu, L., Kittur, A. and Kraut, R. (2014) Searching for analogical ideas with crowds. *CHI 2014*, April 26 - May 01 2014, Toronto, Canada.

Yu, L., Nickerson, J. and Sakamoto, Y. (2012) Collective creativity: where we are and where we might go. *Collective Intelligence 2012*, Boston, USA. 18.-20.4.2012.

*Zembylas, M. and Vrasidas, C. (2005) Globalization, Information, and Communication Technologies, and the Prospect of a ‘Global Village’: Promises of Inclusion or Electronic Colonization?. *Journal of Curriculum Studies*, 2005, 37, 1: 65-83.

*Zettsu, K. and Kiyoki, Y. (2006) Towards knowledge management based on harnessing collective intelligence on the web. *Lecture Notes in Artificial Intelligence*, 4248, 350-357.

Zhao, Y. and Zhu, Q. (2012) Evaluation on crowdsourcing research: Current status and future direction. I*nf Syst Front.*

Zott, C., Amit, R. and Massa, L. (2011) The business model: Recent developments and future research. *Journal of Management,* 37, 1019-1042.

Ågerfalk, P. and Fitzgerald, B. (2008) Outsourcing to an unknown workforce: Exploring opensourcing as a global sourcing strategy. *MIS Quarterly*, 32, 2.

Appendix A: Case study protocol

This study focuses on collective intelligence and its emergence on crowdsourcing sites. It has been claimed that crowdsourcing facilitates, uses, or benefits from collective intelligence, but instead of thorough analyses, the discussion has been more on the level of metaphors. The goal of this study is to find out, whether crowdsourcing can really be connected to phenomena that can be considered to be collective intelligence. In addition the aim is to increase understanding on the exact mechanisms that lead to emergence of collective intelligence. The research questions this study seeks to answer are 1) how collective intelligence is manifested on crowdsourcing sites? and 2) how important collective intelligence is for the functioning of the crowdsourcing sites?

Using the insights from the extant literature it is now possible to construct a theoretical framework to guide the research. As discussed above, collective intelligence refers to phenomena, where the intelligence of a group can be considered to be at least partially independent and usually greater than the intelligence of individuals forming the group. For the purposes of the study the framework from the literature review was simplified according to figure 1. It is assumed that by collecting data about the elements presented in the figure all the relevant aspects will be covered and an understanding on collective intelligence at the site can be developed.

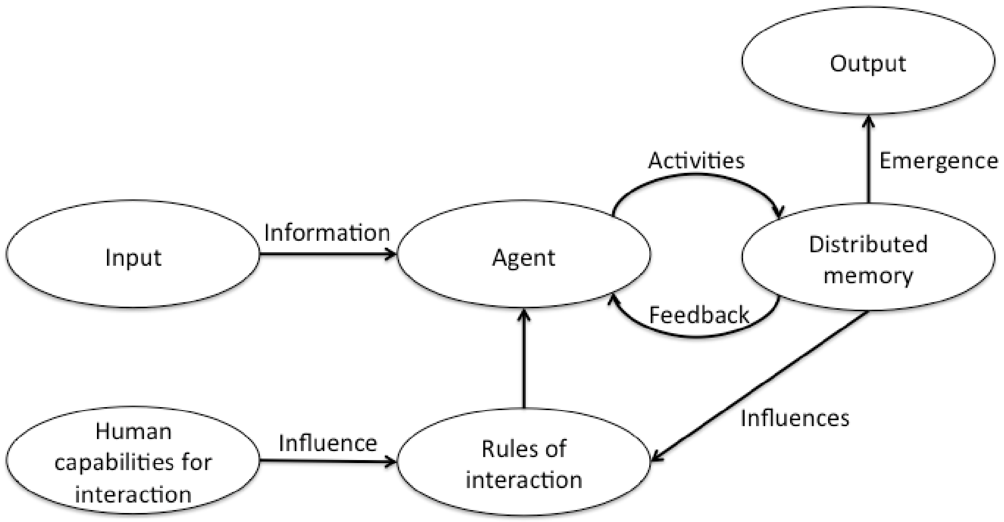


Figure 0.1: The theoretical framework of collective intelligence used to guide data collection.

According to the framework, the human capabilities for interaction, such as intelligence, trust, motivation and other psychological and cultural factors, together with environmental constraints, create the rules of interaction. Inputs to the system arrive through cognitive agents. An agent processes and integrates information from the outside and feedback from the distributed memory and performs actions according to more or less strict rules. The distributed memory is the shared environment of the agents, which stores the information they create. Actions can also change the state of distributed memory. Changes to memory are fed back to the agent, and may also change the environmental constraints. Out of the multiple interactions between agents and distributed memory emerges the output of the system. Agents, their rules of interaction, distributed memory and environmental constraints form a complex adaptive system, which reacts to information from outside. The output is the emergent property of the system and may demonstrate wisdom of the crowds: the decisions made by the system as a whole may be of better quality than individuals are capable to produce alone. These high quality decisions result from diversity, independence and information aggregation.

Data collection procedures

The procedures described in this case study protocol are used to guide the data collection for the purposes of the study. This study explores three cases of crowdsourcing in innovation and product development context: OpenIDEO, Quirky and Threadless. Data is collected as participant-observer on the selected sites. As a participant, each site is visited on at least 30 days. Encountered web pages relevant to the investigation are collected using Evernote and Evernote Web Clipper software. The actions and experiences during the observation are documented in a research diary on Evernote software. All collected data and diary entries are loaded to Dedoose qualitative analysis software and coded using the coding scheme derived from conceptual frameworks of collective intelligence, innovation processes and crowdsourcing. Each case study is documented in a case description featuring the following topics:

1. Rules of interaction
2. Activities
3. Feedback
4. Agents

Case descriptions are organized according to innovation processes used on the case sites.

Appendix B: Coding scheme

Code	Definition	Example	OpenIDEO	Quirky	Threadless
Collective intelligence	References to collective intelligence	DISCARDED	9	0	0
Adaptability	Changing one's structure to fit the environment: individuals, rules or the system.	DISCARDED	0	0	0
Robustness	Even if some parts fail, the system stays functional	DISCARDED	0	0	0
Redundancy	The same information represented in many places	DISCARDED	0	0	0
Emergence	Rise of system level properties that are not present in its components	DISCARDED	2	1	0
Randomness	Elements of randomness in the system	DISCARDED	6	1	3
Human capabilities	The factors affecting a person's ability to interact with other human beings	DISCARDED	5	1	1
Input	Inputs to the system. All the information the agents have access to.	DISCARDED	170	80	28
Output	Outputs of the system. Descriptions of results.	The results of this challenge will be presented at the Digital Agenda Assembly in Brussels on June 21st and 22nd, and the European Commission is committed to implementing some of the top concepts thereafter.	61	31	9
Agents	Descriptions of users and their characteristics.	My background is in commercial real estate development so the idea of rethinking and repurposing space for community vibrancy really resonates with me.	65	51	79
Distributed memory	Information storage shared between the agents.	DISCARDED	4	0	0
Feedback	Feedback to users from the system or from each other. Descriptions of feedback functionality.	The amount of feedback and collaboration you get is overwhelming. From all over the world, in different time zones people have commented on my concepts, and everyone brings a new view to the table – from their part of the world and their background.	1112	1038	319
Interaction	Interactions and activities taken by the users. Descriptions of what users actually do.	We started out by talking with everybody we could – architects, investors, the planning commission, local community members, and others – to get a sense of what was appealing to them, what they saw as	270	120	72

		roadblocks to success, and what they needed from us in order to get on board with our efforts.			
Rules	Explicit and implicit rules of interaction. Descriptions of what is considered correct behavior.	Stay Optimistic, Positive and Respectful	314	396	299
Crowdsourcing	Outsourcing the tasks traditionally performed by an organization to an undefined crowd, usually through an open call posted to Internet.	DISCARDED	7	1	0
Business model	Descriptions or references to ways how the site makes money.	There is a business model: OpenIDEO facilitates innovation process, the sponsor pays the costs and community does the work	49	116	84
CS process	Descriptions of how interaction between the site and users proceeds. Different phases of activity.	And stay tuned: in the next few weeks we'll be launching a new challenge phase called Realisation, which will enable the students of 100K Cheeks to share their implementation progress with the entire OpenIDEO community.	83	100	26
Community	Descriptions of community of users related to the site.	The second thing OpenIDEO offered me was an opportunity to be part of an open source community. I am fascinated by the open source concept and by how people love to collaborate and share passions online.	289	229	365
Gamification	Game-like elements on the site or user interface	Translators can be rewarded with OpenIDEO Badges.	33	50	112
Learning	References to learning new skills or knowledge.	And if you're thinking about setting up a social enterprise or are in your early stages of one – catch some tips on Visualising Your Business Model from OpenIDEO's Tom Hulme.	102	40	81
Marketing	Descriptions of marketing efforts and references to elements on the site that support marketing.	On OpenIDEO we celebrate that our community members can join our challenges in whatever way works best for them: from adding content and comments, to reading posts and getting inspired.	272	205	182
Motivation	Factors that motivate or are assumed to motivate participation. Descriptions of why do they participate.	One week to go guys – get your ideas posted to help us re-imagine the future of food. You might even win the chance to join us in sunny Queensland at the IDEAS 2011 Festival. And check out IDEO's Paul Bennett talking about the challenge and his vinyl record obsession.	48	25	26
Platform	Descriptions of website and user interface and it's functionality.	Collaboration map. This somehow tracks how the concepts are build: what are the parts. Might be possible to evaluate whether it is more than the sum of the parts...	98	179	121

Success factor	Descriptions of best practices and features that are considered to contribute to success.	DISCARDED	2	10	3
Tasks	Descriptions of tasks the site asks users to perform, either explicitly or implicitly.	We'd love you to share any examples you've seen of new and inspiring ways to develop soft or hard skills that are happening beyond the classroom.	503	364	261
User experience	Personal experiences from using the site.	Found the missions on the left panel of Inspirations site. Still don't really understand them. How do they differ from Themes?	183	197	203
Wisdom of crowds	Phenomenon where, under certain conditions, aggregated estimate of a large and diverse group may be more accurate than the estimates of any single individual in the group.	DISCARDED	0	3	1
Aggregation	The combination of individual pieces of information to form a synthesis or collective estimation	DISCARDED	1	27	21
Bias	Evidence of the tendency of individuals and groups to make systematic errors in decision-making situations	I noticed I decide whether to open a concept from a list view at least partly based on the applause it has already gathered.	24	69	81
Decision making	References to the process of making decisions, both individually and in groups	Eventually a selection of concepts are chosen as winners.	50	126	44
Independence	The decision of an individual is not influenced by the decisions of other individuals	DISCARDED	11	37	20
Diversity	Descriptions of diversity of users and impacts of it.	From all over the world, in different time zones people have commented on my concepts, and everyone brings a new view to the table – from their part of the world and their background.	8	4	1
Innovation process	Descriptions of underlying innovation process.	Process description with the current phase and numerical measurements is clear, bright and colorful and immediately noticeable on the top of the page. Gives a lot of information to the user, fast and easy.	86	177	71
Problem definition	References to problem definition phase of innovation process.	How might we, for instance, help startups access funding across stages of development? Or help them find resources when working across countries? Or	755	368	48

		foster a culture of experimentation?			
Idea generation	References to idea generation phase of innovation process.	It all starts with a good idea. After all, a good idea attracts a lot of supporters and is easier to make happen. Finding that good idea, however, is the challenging part!	567	289	165
Idea evaluation	References to idea evaluation phase of innovation process.	The evaluation phase allowed everyone to have their say on which concept should go forwards to become the OpenIDEO logo. Set criteria were used to make this judgement; things like fit with our community principles, and just how much they loved it.	273	949	377
Development	References to development phase of innovation process.	As I mentioned, getting the Grand Rapids community stakeholders onboard has been hugely important. Also, being open to prototyping – and potentially failing in the process – has been big for us.	532	444	172
Implementation	References to implementation phase of innovation process.	Comprised of refurbished shipping containers, Intermodal will house local food producers, artists, or other merchants to showcase their products and connect locally with consumers.	152	89	17

Appendix C: Wisdom of crowds at OpenIDEO

Summary of the dataset

```
## E-waste challenge data consists of 106 observations of 14 variables
.
##  Unemployment challenge data consists of 149 observations of 14 var
iables.
##  Business celebration challenge data consists of 95 observations of
14 variables.
```

A sample of data:

Order	Concept	Views	Comments	Applause	Shortlist
1	concept_73	596	19	11	Rejected
2	concept_60	696	26	25	Rejected
3	concept_58	225	3	7	Rejected
4	concept_35	394	28	36	Shortlisted
5	concept_86	309	15	15	Shortlisted
6	concept_106	242	15	11	Rejected

The most interesting thing to explore is the relationship between different variables and selection to shortlist, as this offers an opportunity to compare preferences of the crowd to expert decision.

Descriptive statistics

	Views	Comments	Applause
nbr.val	350.000	350.000	350.000
nbr.null	0.000	35.000	1.000
nbr.na	0.000	0.000	0.000
min	13.000	0.000	0.000
max	1448.000	49.000	43.000
range	1435.000	49.000	43.000
sum	83524.000	2216.000	2313.000
median	185.000	4.000	5.000
mean	238.640	6.331	6.609
SE.mean	10.161	0.354	0.323
CI.mean.0.95	19.985	0.696	0.635
var	36138.334	43.827	36.514
std.dev	190.101	6.620	6.043
coef.var	0.797	1.046	0.914

```
##
##    Rejected Shortlisted
##         290          60

## 17 % of concepts have been selected on the shortlist and 4 % of con
cepts are winners.
```

Concepts have much more views than comments or applause, which makes sense as it is easier just to view a concept than do something about it. Interestingly statistics on comments and applause are very similar to each other. They have similar ranges, means, medians, standard deviations, sums and even skewness. 20 concepts per challenge have been selected on the shortlist. Acceptance rate is higher than on either Threadless or Quirky.

Descriptive statistics by challenge

```
## openideo$Challenge: celebrate
##                      Views   Comments   Applause
## median       9.100000e+01  3.0000000  3.0000000
## mean         1.559684e+02  4.9157895  4.9789474
## SE.mean      1.572530e+01  0.5307273  0.4295832
## CI.mean.0.95 3.122296e+01  1.0537715  0.8529476
## var          2.349209e+04 26.7587906 17.5314670
## std.dev      1.532713e+02  5.1728900  4.1870595
## coef.var     9.827073e-01  1.0523010  0.8409527
## ------------------------------------------------------------
## openideo$Challenge: ewaste
##                      Views   Comments   Applause
## median         166.500000  6.0000000  5.0000000
## mean           233.094340  7.1886792  6.7924528
## SE.mean         17.608514  0.7243813  0.6285677
## CI.mean.0.95    34.914430  1.4363144  1.2463337
## var          32866.333872 55.6212040 41.8803235
## std.dev        181.290744  7.4579625  6.4715009
## coef.var         0.777757  1.0374593  0.9527487
## ------------------------------------------------------------
## openideo$Challenge: unemployment
##                      Views   Comments   Applause
## median         235.000000  5.0000000  5.0000000
## mean           295.295302  6.6241611  7.5167785
## SE.mean         16.248801  0.5495143  0.5353124
## CI.mean.0.95    32.109621  1.0859075  1.0578430
## var          39339.506802 44.9929258 42.6973517
## std.dev        198.341894  6.7076766  6.5343211
## coef.var         0.671673  1.0126077  0.8692981
```

There is some variability between challenges. Celebrating innovative businesses challenge appears to have been the least popular of the three. That challenge is also missing data on what happened after the shortlist selection. This is not a problem, though, as here the focus is on events before the shortlist selection.

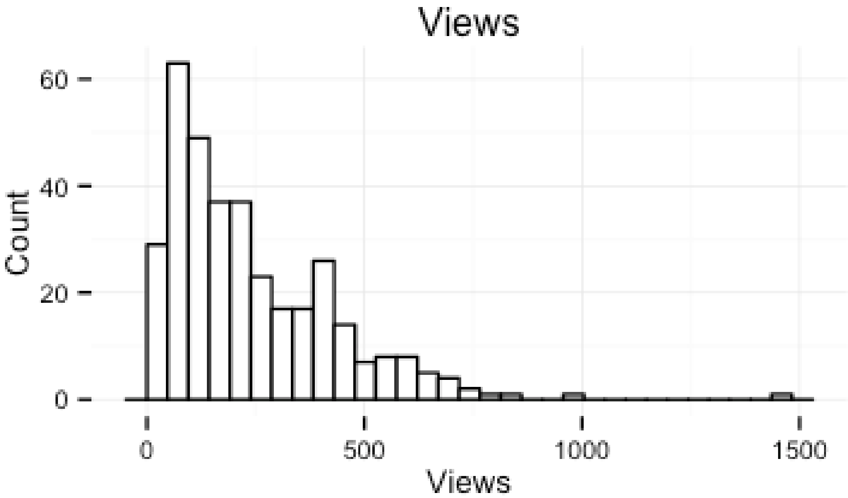
Descriptive statistics by shortlist status

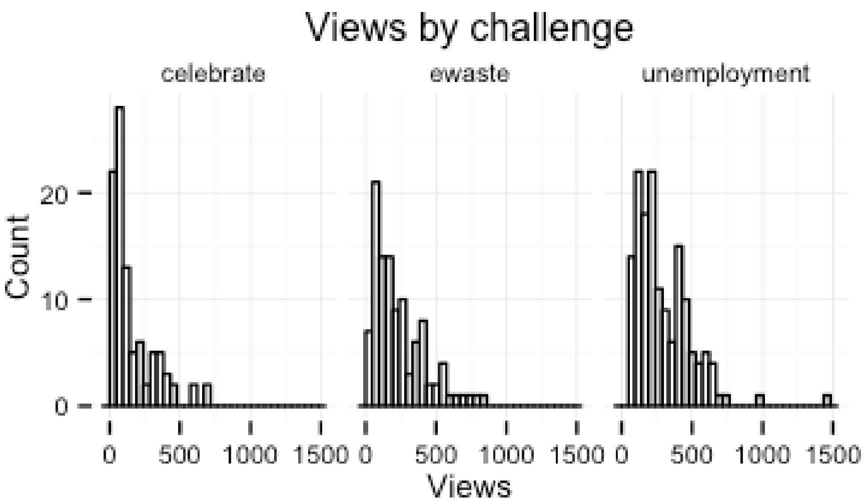
```
## openideo$Shortlist: Rejected
##                     Views   Comments   Applause
## median         151.500000  4.0000000  4.0000000
## mean           207.862069  5.2758621  5.6034483
## SE.mean          9.931698  0.3204059  0.2960060
## CI.mean.0.95    19.547631  0.6306250  0.5826010
## var          28605.198902 29.7713877 25.4096766
## std.dev        169.130715  5.4563163  5.0408012
## coef.var         0.813668  1.0342037  0.8995891
## --------------------------------------------------------
## openideo$Shortlist: Shortlisted
##                     Views   Comments  Applause
## median       3.710000e+02 10.5000000  9.000000
## mean         3.874000e+02 11.4333333 11.466667
## SE.mean      2.783574e+01  1.1652615  1.021096
## CI.mean.0.95 5.569918e+01  2.3316828  2.043208
## var          4.648970e+04 81.4700565 62.558192
## std.dev      2.156147e+02  9.0260765  7.909374
## coef.var     5.565687e-01  0.7894528  0.689771
```

On average the shortlisted designs gather about the double the amount of views, comments and applause compared to rejected designs. There is also more variance in these statistics among the shortlisted designs than among the rejected designs.

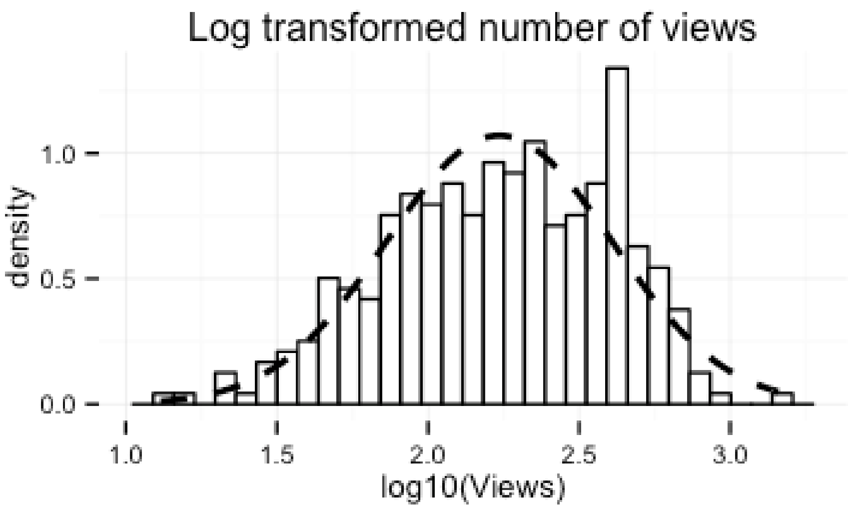
Distributions of variables

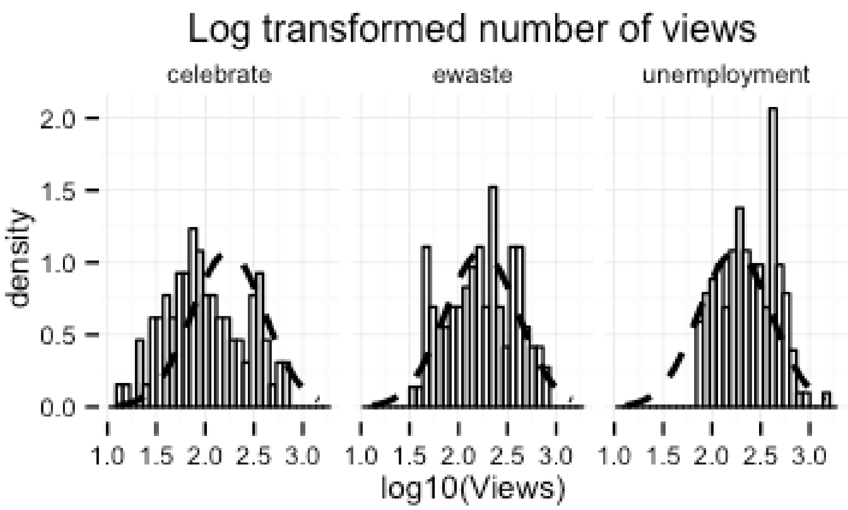
Views





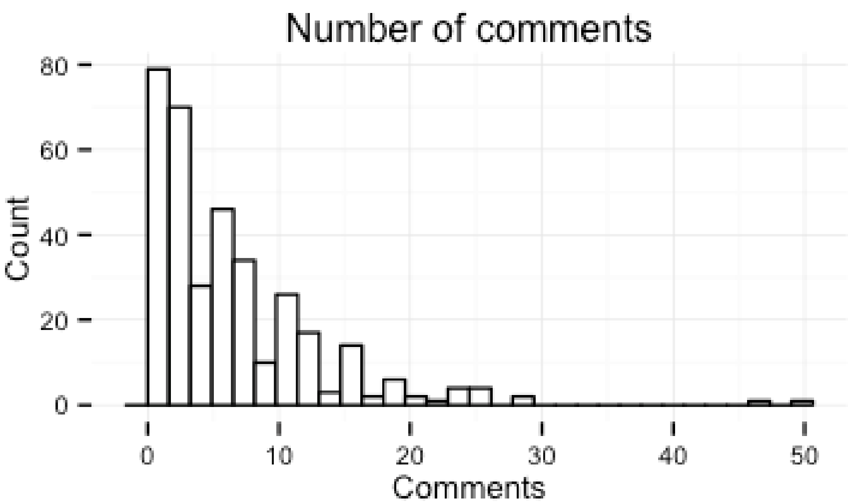
The distribution of views is right-skewed due to natural limit at zero views and no limit at the other end of the scale. The situation is the same both at the aggregate level and within individual challenges.

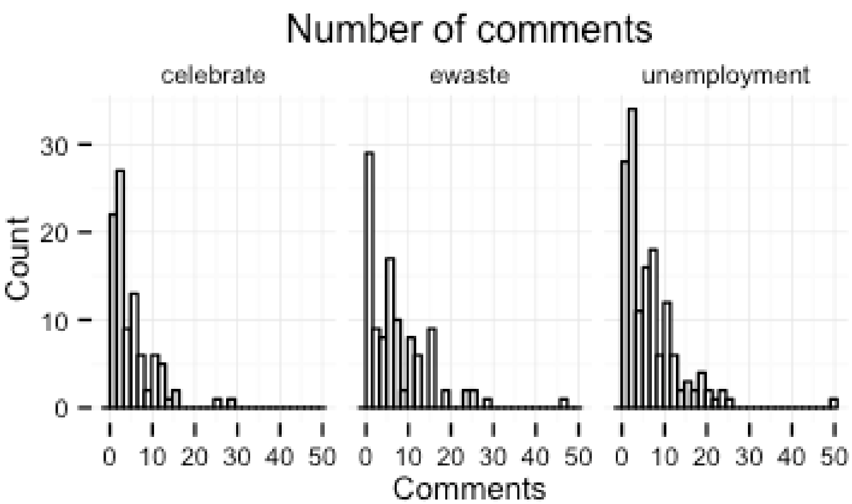




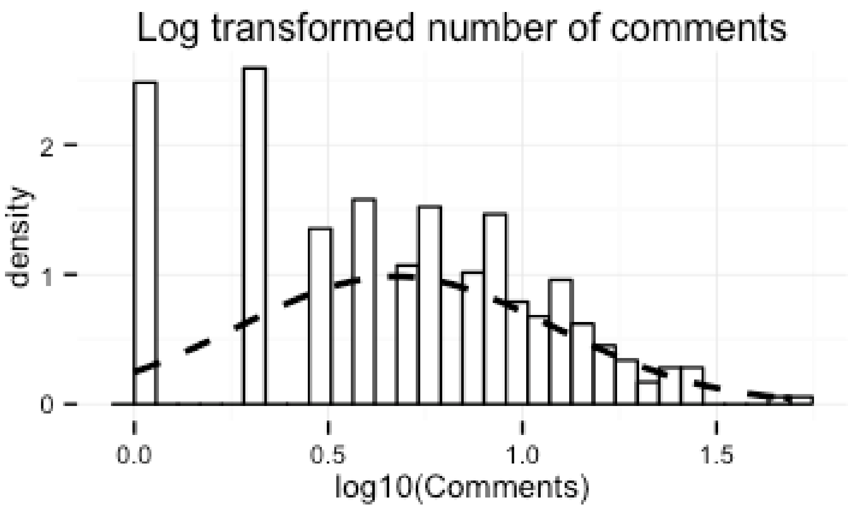
Log-transforming the number of views results in a near-normal distribution for the combined data set, but not for the individual challenges, perhaps due to smaller amount of data.

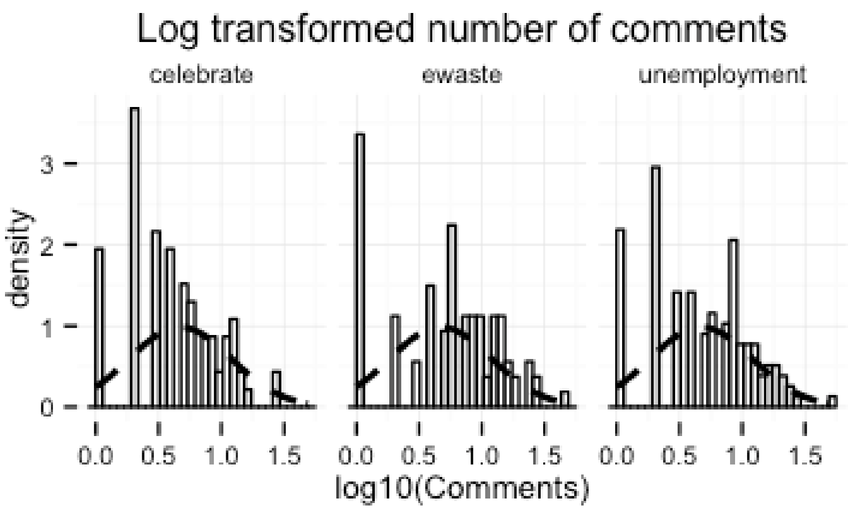
Comments





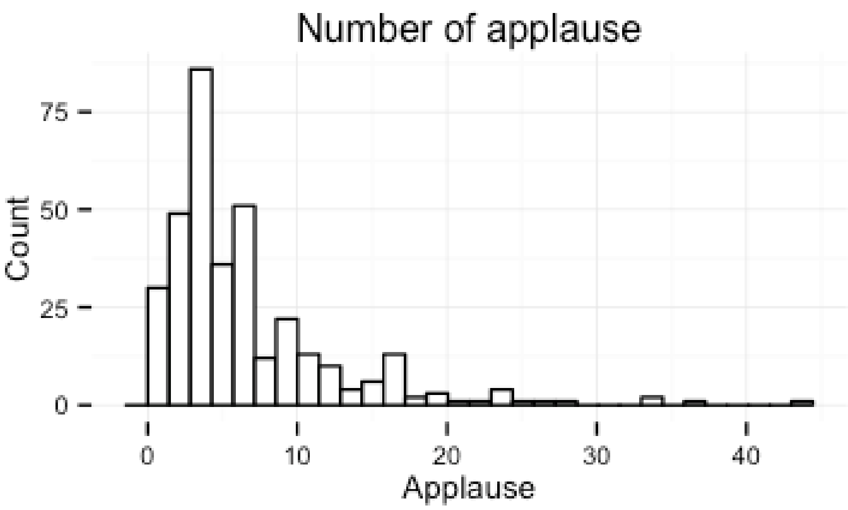
Situation is the same with the distribution of number of comments as with the number of views.

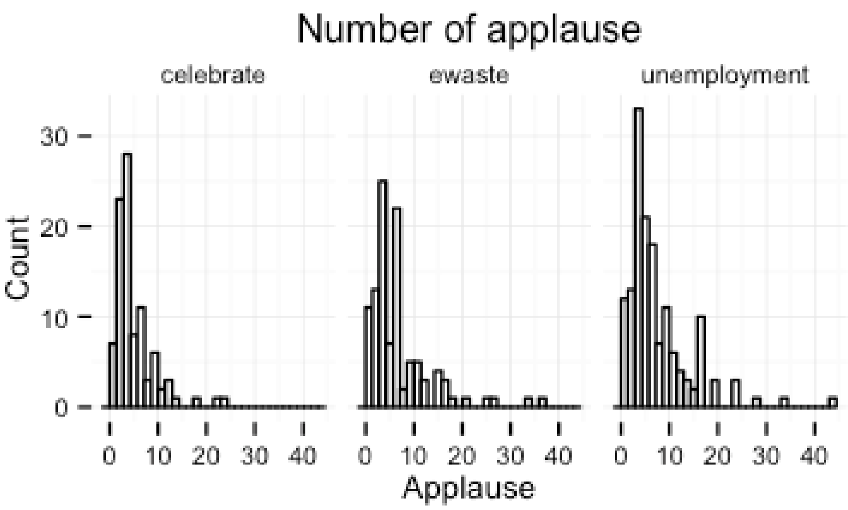




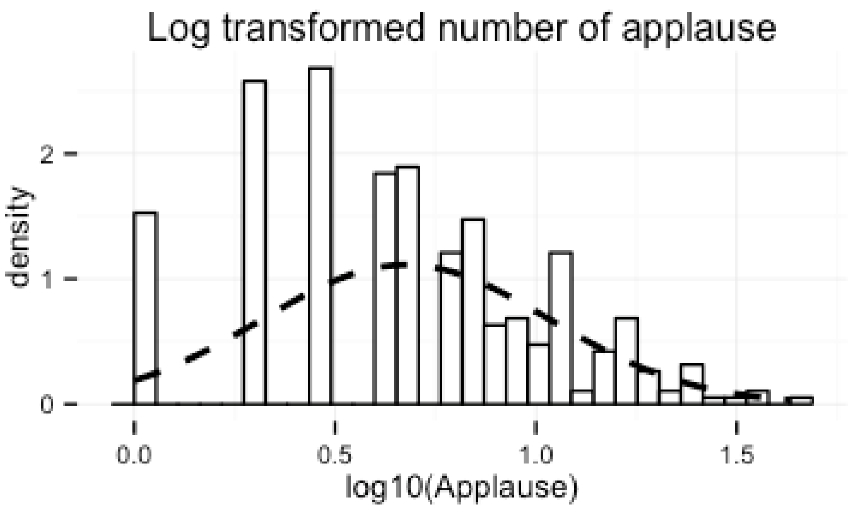
Log-transformation does not result in normal-looking distribution. This might be problematic in further analysis.

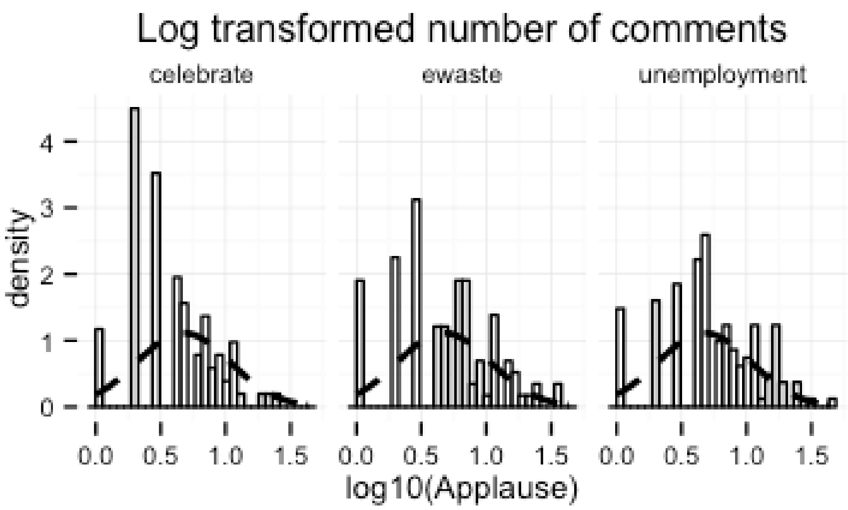
Applause





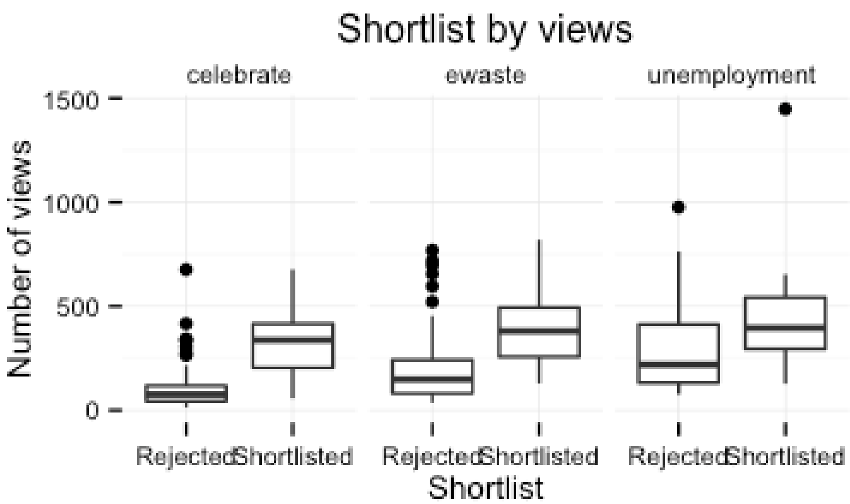
Distribution of number of applause repeats the familiar shape. In contrast to capped scores used at Threadless the measurements used at OpenIDEO are not limited, which leads to less helpful distributions.

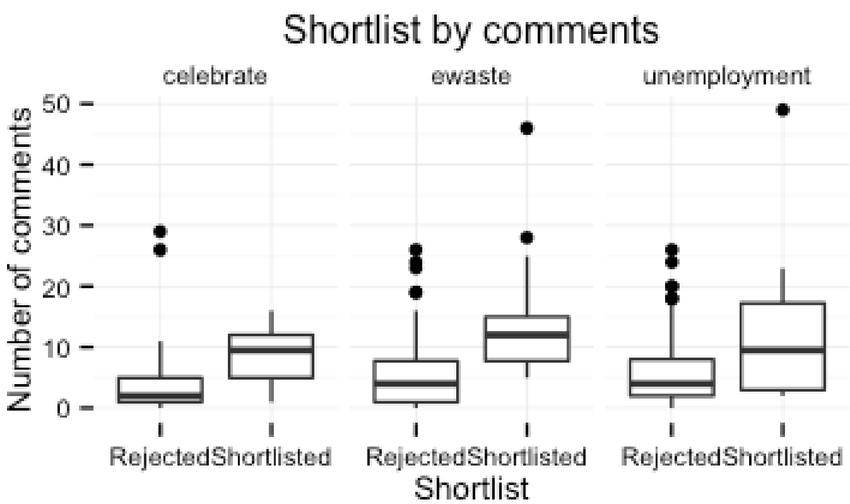


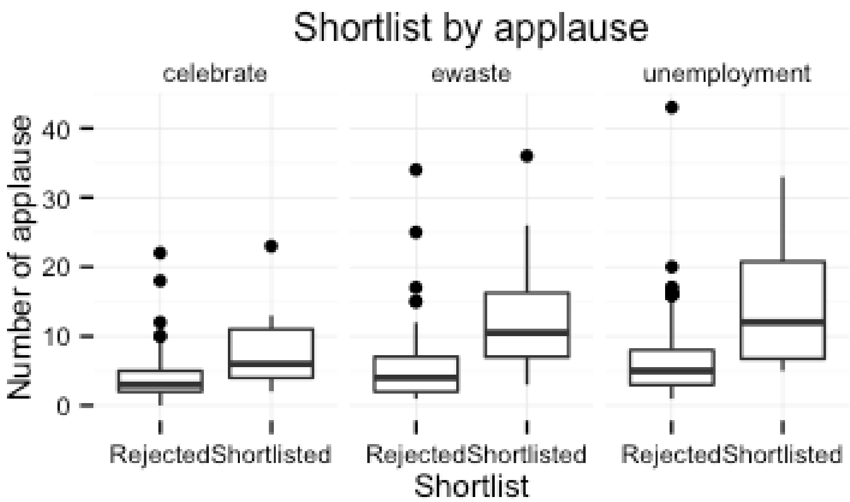


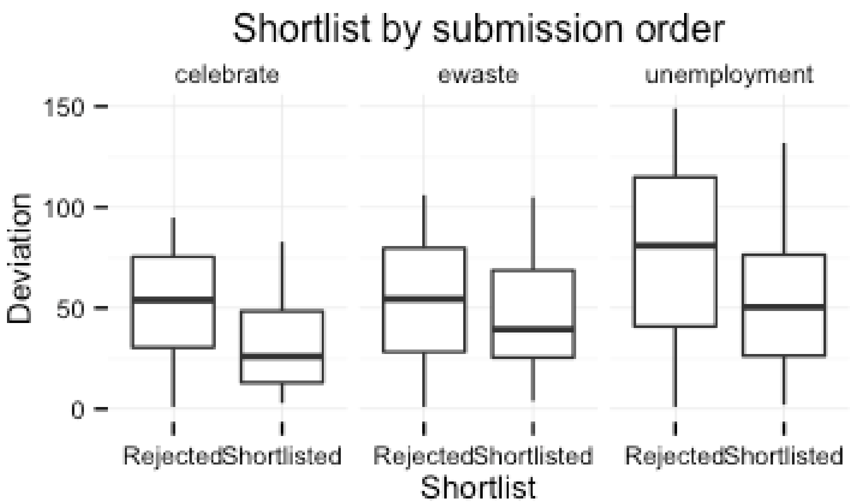
Log-transformation does not result in normal distribution. This might be challenging in further analysis.

Relationships between variables









In all challenges the shortlisted designs tend to have more views, comments and applause. There also appears to be a small tendency for shortlisted designs to have been submitted earlier in the challenge. Unfortunately data on exact submission dates is not available in the dataset.

Logistic regression analysis

```
## glm(formula = Shortlist ~ ApplauseStd + CommentsStd + ViewsStd +
##     OrderStd, family = "binomial", data = train)
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.9707060	0.2268870	-8.685850	0.0000000
ApplauseStd	0.2982571	0.2309490	1.291441	0.1965507
CommentsStd	0.2860971	0.2352681	1.216047	0.2239671
ViewsStd	0.8870787	0.2344347	3.783906	0.0001544
OrderStd	-0.0004701	0.2238484	-0.002100	0.9983245

When trying to predict the shortlist status based on the standardized number of views, comments, applause and submission order, only the number of views is statistically significant. Therefore, a model containing only the number of views is used.

Correlations between variables:

	Order	ViewsStd	CommentsStd	ApplauseStd
Order	1.0000000	-0.3576367	-0.3445924	-0.3486417
ViewsStd	-0.3576367	1.0000000	0.6655894	0.6288069
CommentsStd	-0.3445924	0.6655894	1.0000000	0.6653346
ApplauseStd	-0.3486417	0.6288069	0.6653346	1.0000000

There is moderate correlation between variables, further suggesting the most of the information available is already contained in the most significant variable.

```
## glm(formula = Shortlist ~ ViewsStd, family = "binomial", data = tra
in)
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.920548	0.2165000	-8.870892	0
ViewsStd	1.187853	0.2031993	5.845752	0

In the simpler model the number of views is statistically significant predictor of shortlist status.

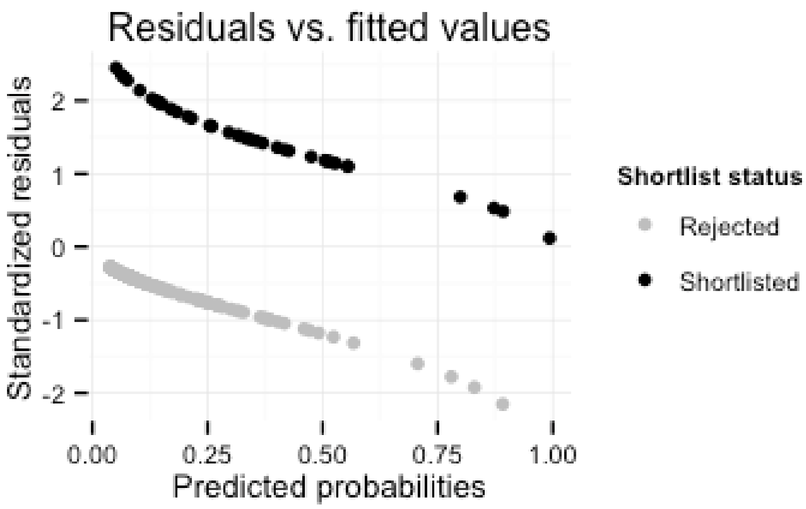
Resid. Df	Resid. Dev	Df	Deviance	Pr(>Chi)
239	168.2785	NA	NA	NA
242	173.7927	-3	-5.51421	0.1377912

The more complex model has slightly smaller residual deviance than the simple model, but because other variables were far from being statistically significant, there is a good change that the more complex model is just fitting the noise. The performance of models is not very different overall.

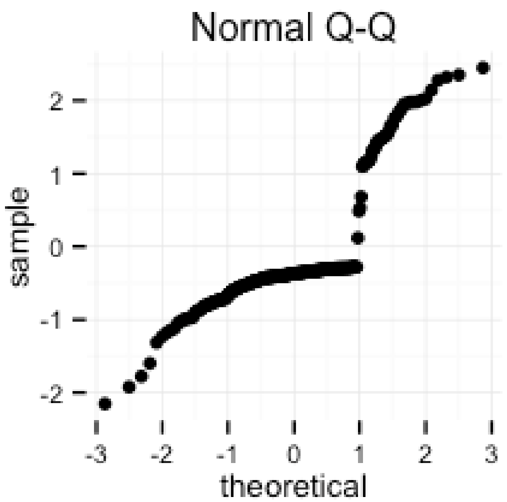
```
## modelChi       47.15083
## chidf          1
## chisq.prob     6.572853e-12
## Pseudo R^2 for logistic regression
## Hosmer and Lemeshow R^2    0.213
## Cox and Snell R^2          0.176
## Nagelkerke R^2             0.295
## Odds ratios:
## (Intercept)    ViewsStd
##       0.147       3.280
## Confidence intervals:
##             0.5 % 99.5 %
## (Intercept) 0.080  0.246
## ViewsStd    2.017  5.792
```

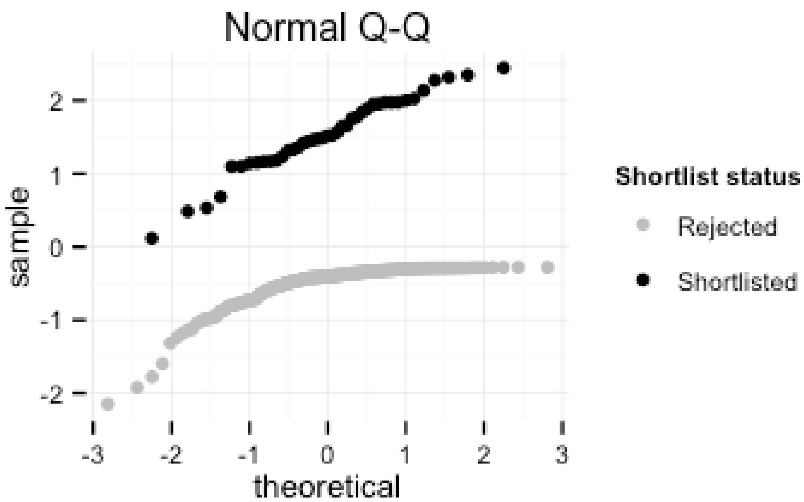
The model is clearly better than random baseline model, but the effect size is only moderate, as estimated by the pseudo R^2 statistics. On average, a concept gaining one standard deviation more views in a challenge increases the odds of that concept being selected on the shortlist by a factor of about 3.

Model diagnostics



The model makes largest mistakes by missing the concepts that get on the shortlist, similarly to problems with Threadless model. Here the effect is smaller though, presumably due to larger ratio of submissions being selected.





QQ-plots show the same issue. Because selected concepts are relatively rare, the model tends to predict that nothing is selected, but suffers only small punishments for the mistakes it makes. The predictive performance might be almost trivial.

```
## 8 observations have residuals larger than 2 standard deviations.
##  23 observations have leverage more than 2 times larger than the av
erage.
```

Observations with large residuals:

```
##
##          Rejected Shortlisted
##   FALSE      202          34
##   TRUE         1           7
```

Most of the observations with large residuals are shortlisted concepts.

Observations with large leverage:

```
##
##          Rejected Shortlisted
##   FALSE      193          28
##   TRUE        10          13
```

With large leverage the pattern is less clear. The shortlisted concepts are still over represented, but not as clearly as with the large residuals.

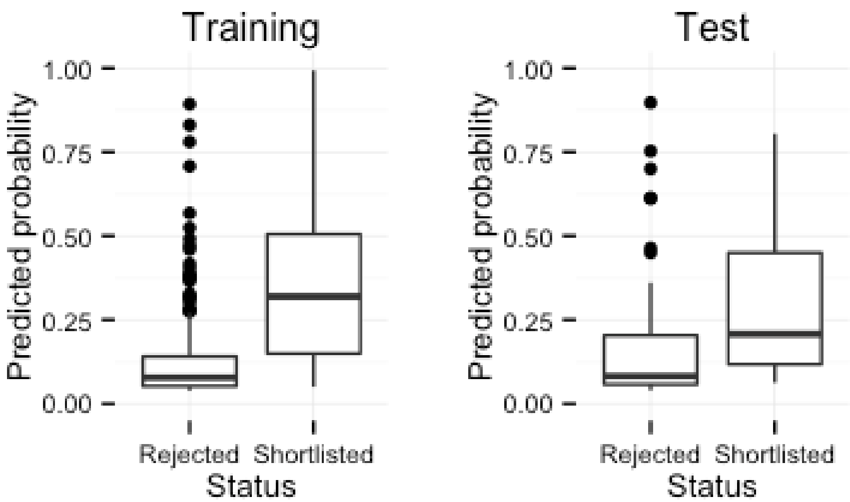
```
## train$Shortlist[train$large.leverage]: Rejected
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.335   1.538   1.770   2.085   2.600   3.393
## --------------------------------------------------------
## train$Shortlist[train$large.leverage]: Shortlisted
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.338   1.632   1.693   1.971   1.803   3.399

## train$Shortlist[!train$large.leverage]: Rejected
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -1.11100 -0.83490 -0.50870 -0.35810 -0.04465  1.28400
## --------------------------------------------------------
## train$Shortlist[!train$large.leverage]: Shortlisted
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.84850  0.09187  0.41950  0.58540  0.99190  5.81200
```

It is not obvious what causes some observations to have a large leverage.

The concepts with least views are rarely selected on the shortlist, but after that the pattern is not clear. Concepts with more than average number of views seem to have quite equal changes of getting on the shortlist. This figure also shows the small trend of older concepts having slightly more views and being selected more often on the shortlist.

Validation

To validate the model the predicted probabilities between training and test sets are compared.



The shortlisted concepts in the test set tend to have higher predicted probabilities than rejected designs, but the model is not accurate enough to discriminate the concepts. Mean squared error for the training set is 0.112 and MSE for the test set is 0.143. The difference is small, which indicates the model fits the training and tests sets well.

Final model

Logistic regression model using number of views to predict shortlist status is fitted using the full data set.

```
## glm(formula = Shortlist ~ ViewsStd, family = "binomial", data = ope
nideo)
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.846860	0.1717637	-10.752329	0
ViewsStd	1.016392	0.1550435	6.555527	0

```
## modelChi       53.53801
## chidf          1
## chisq.prob     2.53686e-13
## Pseudo R^2 for logistic regression
## Hosmer and Lemeshow R^2    0.167
## Cox and Snell R^2          0.142
## Nagelkerke R^2             0.236
## Odds ratios:
## (Intercept)    ViewsStd
##       0.158       2.763
## Confidence intervals:
##             0.5 % 99.5 %
## (Intercept) 0.098  0.239
## ViewsStd    1.890  4.221
```

Appendix D: Wisdom of crowds at Threadless

Summary of the dataset

```
## Dataset consists of 15581 observations of 10 variables.
##  194 rows with missing values have been removed.
```

A sample of data:

	approved_date	design	user	avg.score	score	fives	ones	printed	challenge	printed.binary
44	2012-10-08	design_22499	user_2031	3.30	565	119	52	Not printed	threadless	0
50	2012-10-09	design_19072	user_4755	2.83	738	87	110	Not printed	threadless	0
52	2012-10-10	design_25317	user_2373	2.65	310	18	56	Not printed	threadless	0
53	2012-10-10	design_18499	user_12239	2.36	328	23	85	Not printed	threadless	0
63	2012-10-11	design_27328	user_4861	2.12	186	11	66	Not printed	threadless	0
64	2012-10-11	design_26701	user_2562	2.25	185	10	58	Not printed	threadless	0

Descriptive statistics

	avg.score	score	fives	ones
nbr.val	15581.000	15581.000	15581.000	15581.000
nbr.null	0.000	0.000	1.000	0.000
nbr.na	0.000	0.000	0.000	0.000
min	1.550	102.000	-115.000	13.000
max	4.690	4961.000	4175.000	722.000
range	3.140	4859.000	4290.000	709.000
sum	42904.060	5298739.000	855329.000	1072391.000
median	2.750	294.000	39.000	59.000
mean	2.754	340.077	54.896	68.827
SE.mean	0.003	1.517	0.639	0.312
CI.mean.0.95	0.006	2.973	1.252	0.612
var	0.156	35837.889	6358.350	1519.129
std.dev	0.394	189.309	79.739	38.976
coef.var	0.143	0.557	1.453	0.566

```
##       approved_date       design      user avg.score score fives on
es
## 25681    2013-04-29 design_21031 user_3060      2.92   231  -115
32
##           printed  challenge printed.binary
## 25681 Not printed threadless              0

## 1.32 % of designs have been printed.
```

Average score of designs varies between 1.55 and 4.69, close to smallest and largest possible average scores 1 and 5. Numbers of scores, fives and ones are not limited, and indeed some designs have had thousands of people scoring them. Fives have an outlier: for some reason one of the values is negative. Only a small minority of designs in the data have been printed.

Descriptive statistics by print status

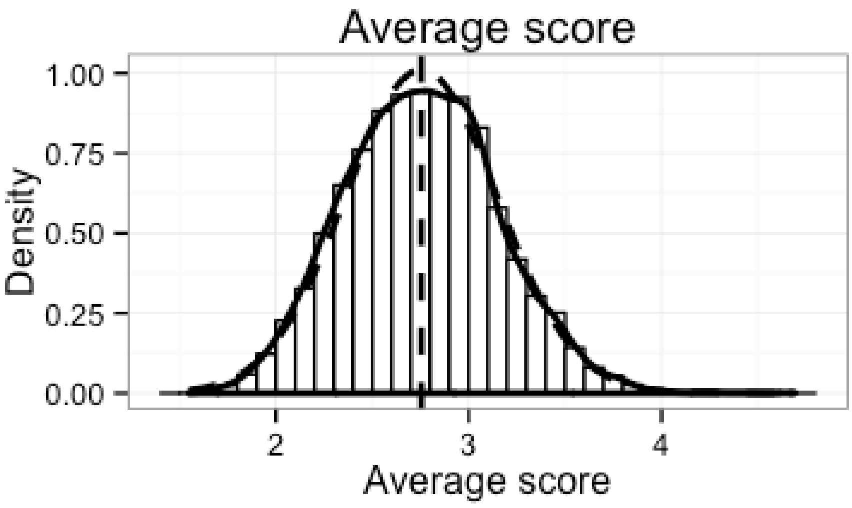
```
## threadless$printed: Not printed
##    avg.score         score            fives              ones
##  Min.   :1.550   Min.   : 102.0   Min.   :-115.00   Min.   : 13.00
##  1st Qu.:2.470   1st Qu.: 221.0   1st Qu.:  23.00   1st Qu.: 44.00
##  Median :2.750   Median : 293.0   Median :  39.00   Median : 59.00
##  Mean   :2.748   Mean   : 338.7   Mean   :  54.07   Mean   : 68.98
##  3rd Qu.:3.010   3rd Qu.: 404.0   3rd Qu.:  65.00   3rd Qu.: 82.00
##  Max.   :4.690   Max.   :4961.0   Max.   :4175.00   Max.   :722.00
## ---------------------------------------------------------
## threadless$printed: Printed
##    avg.score         score            fives             ones
##  Min.   :2.010   Min.   : 104.0   Min.   :   6.0   Min.   : 15.00
##  1st Qu.:2.880   1st Qu.: 268.0   1st Qu.:  46.0   1st Qu.: 38.00
##  Median :3.240   Median : 380.0   Median :  89.0   Median : 53.00
##  Mean   :3.179   Mean   : 439.7   Mean   : 116.9   Mean   : 57.62
##  3rd Qu.:3.500   3rd Qu.: 549.0   3rd Qu.: 151.0   3rd Qu.: 71.00
##  Max.   :4.280   Max.   :1720.0   Max.   :1153.0   Max.   :161.00
```

```
## threadless$printed: Not printed
##                avg.score        score         fives          ones
## median       2.750000000 2.930000e+02   39.0000000   59.0000000
## mean         2.747937045 3.387481e+02   54.0691337   68.9762617
## SE.mean      0.003148342 1.518445e+00    0.6359354    0.3152669
## CI.mean.0.95 0.006171123 2.976332e+00    1.2465086    0.6179604
## var          0.152407815 3.545207e+04 6218.2667649 1528.2699405
## std.dev      0.390394436 1.882872e+02   78.8559875   39.0930933
## coef.var     0.142068188 5.558325e-01    1.4584289    0.5667616
## ---------------------------------------------------------
## threadless$printed: Printed
##               avg.score        score        fives       ones
## median       3.24000000 3.800000e+02 8.900000e+01  53.000000
## mean         3.17941463 4.397463e+02 1.168878e+02  57.619512
## SE.mean      0.03199068 1.637609e+01 7.973699e+00   1.860693
## CI.mean.0.95 0.06307477 3.228810e+01 1.572143e+01   3.668656
## var          0.20979769 5.497615e+04 1.303387e+04 709.746676
## std.dev      0.45803678 2.344699e+02 1.141660e+02  26.641071
## coef.var     0.14406324 5.331936e-01 9.767143e-01   0.462362
```

Printed designs tend to have higher average scores: mean 3.2 for printed desings vs. 2.75 for designs that have not been printed. Numbers of scores and fives are also a little higher compared to designs that have not been printed.

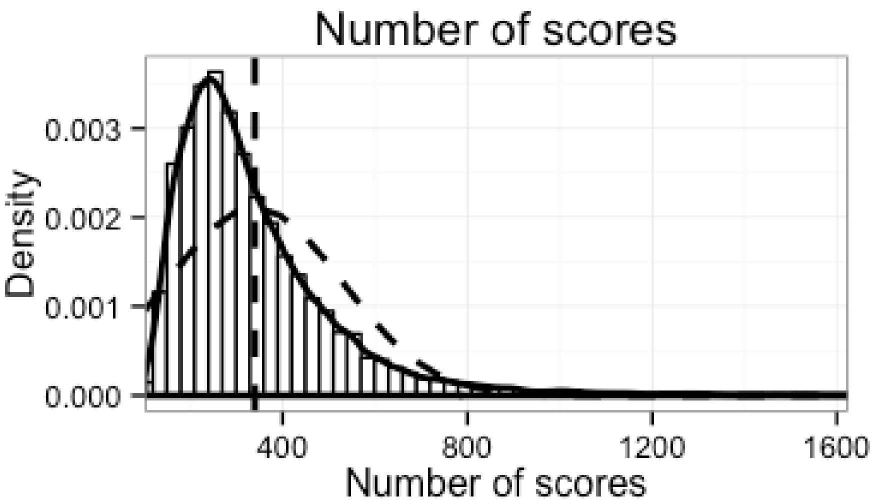
Distributions of single variables

Average score

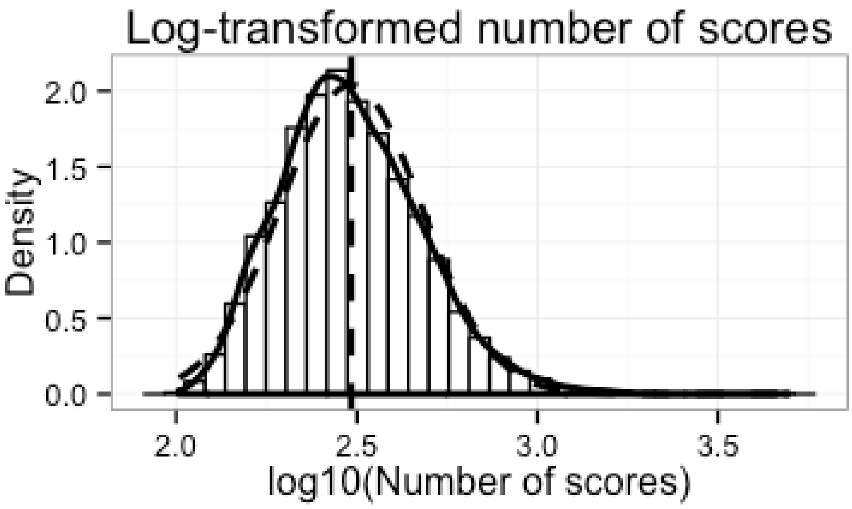


The distribution of average scores (solid black line) appears to follow closely normal distribution with same mean and standard deviations (dashed line).

Number of scores

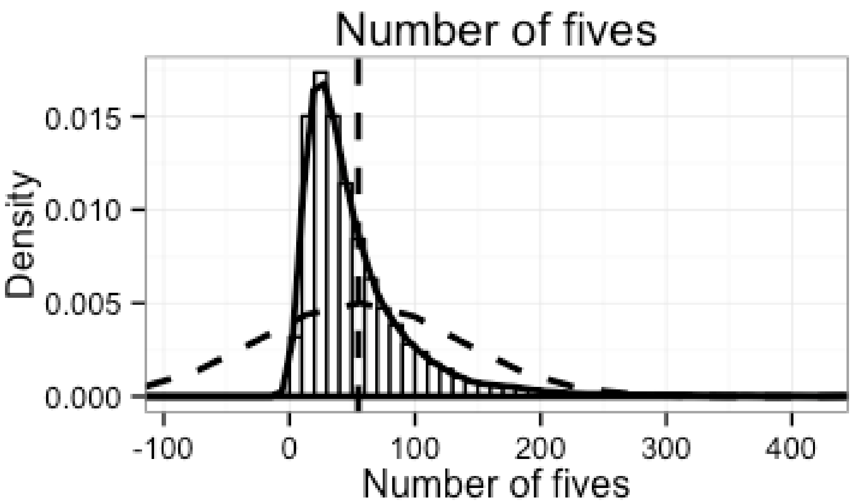


Distribution of number of scores is right-skewed and does not resemble normal distribution (dashed line). Median is a better summary than mean for this distribution.

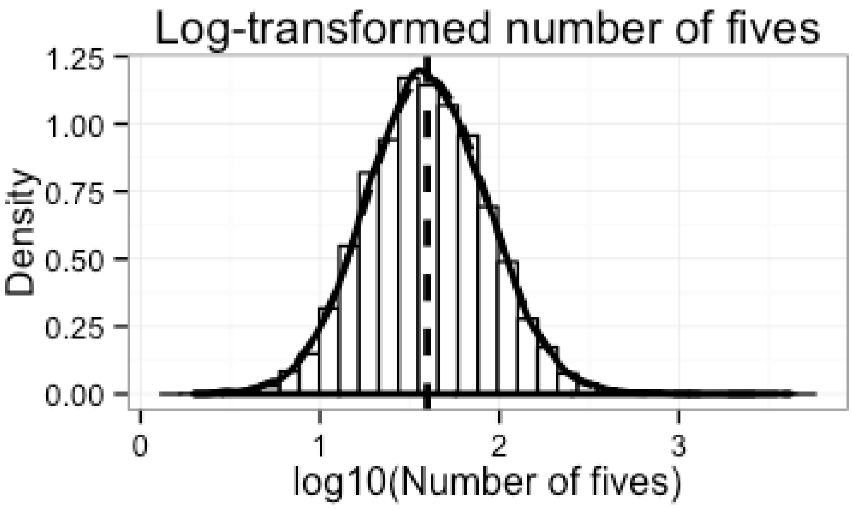


Log transformation fixes the skewness and gets the distribution much closer to the normal distribution.

Number of fives

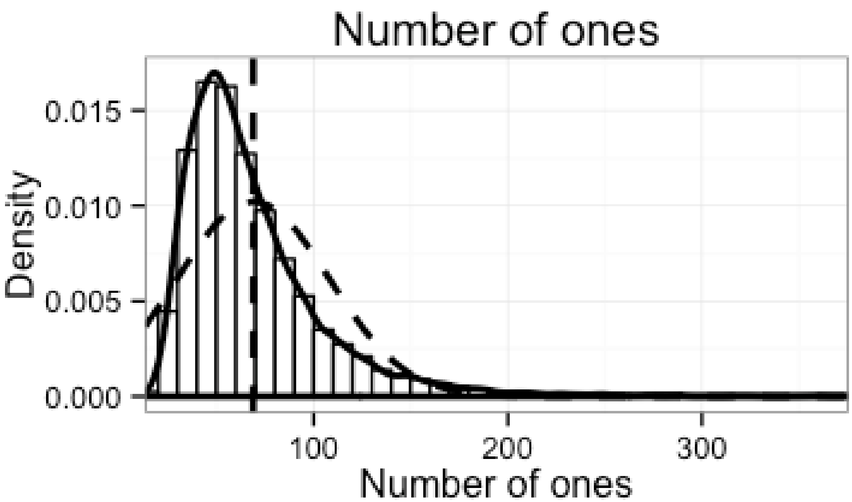


Distribution of number of fives is also right-skewed and does not follow normal distribution.

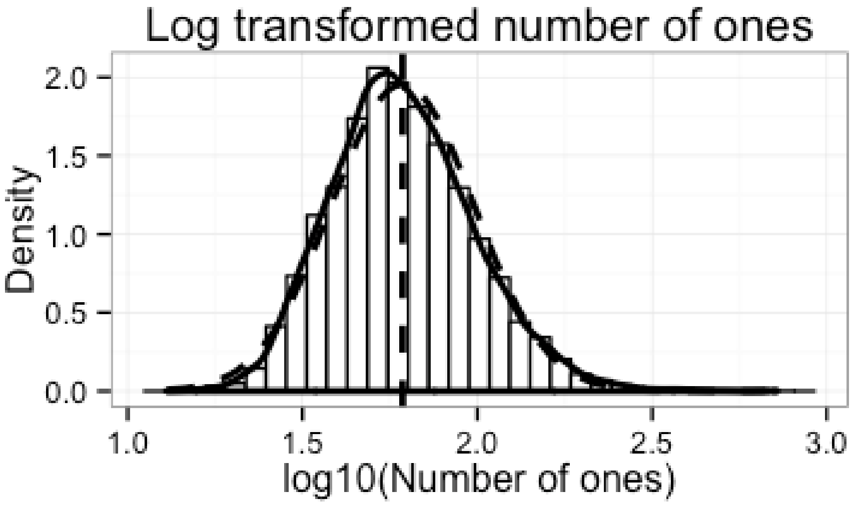


Log transformed data closely resembles a normal distribution

Number of ones

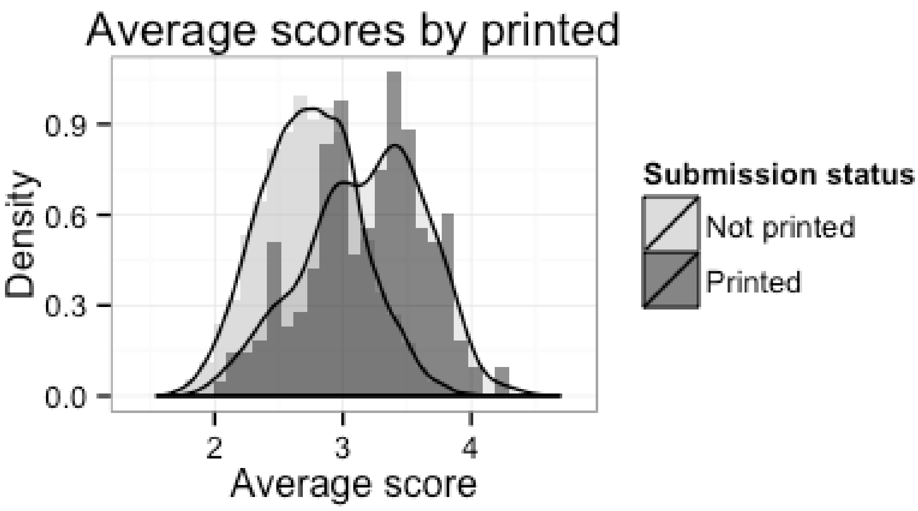


Similar right-skewed distribution again as with the scores and fives.

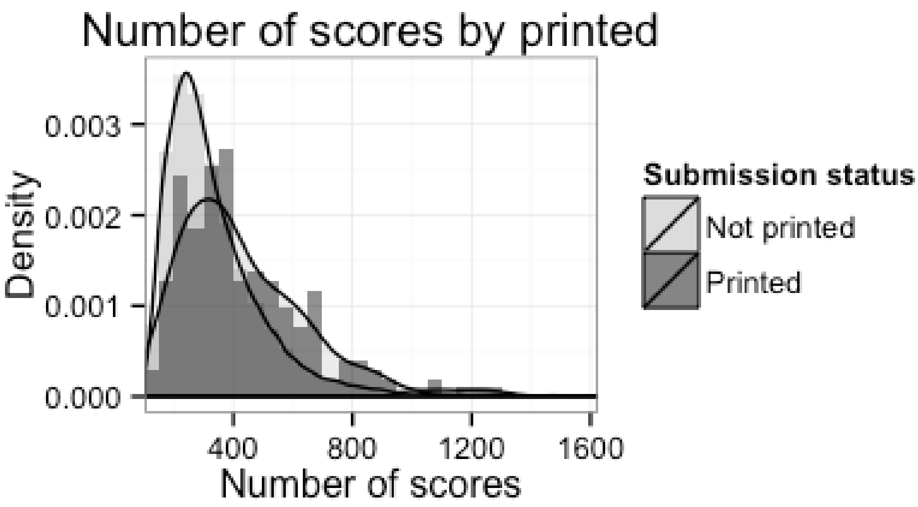


Log transformed data again resembles normal distribution. If scores, fives or ones are used in further analyses, it is probably best to use them in log transformed formats.

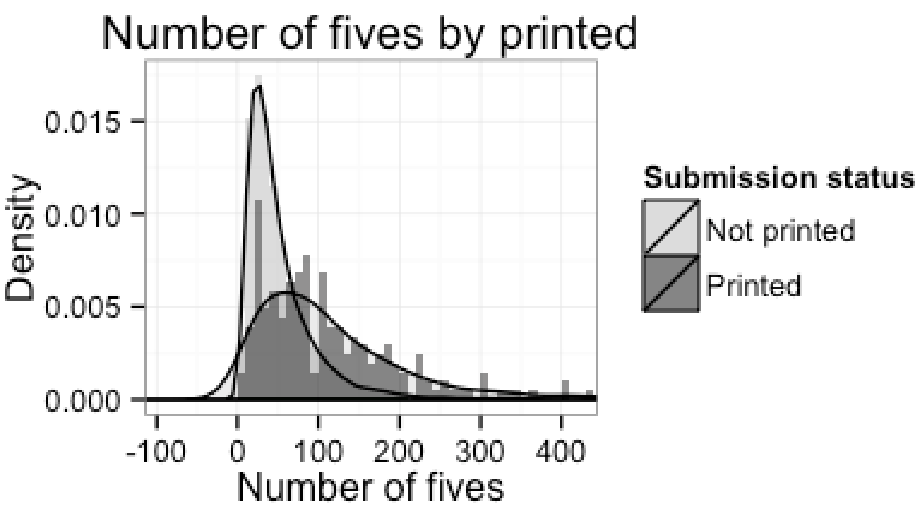
Relationships between variables



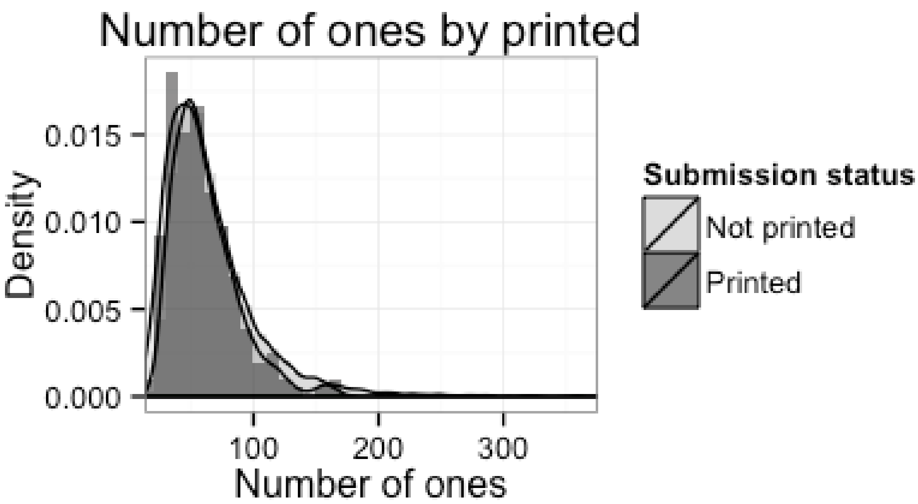
Printed designs have a tendency to have higher average scores than designs that have not been printed, but there's quite a lot of overlap.



Distribution of number of scores is shifted slightly right for printed designs. The difference is small.

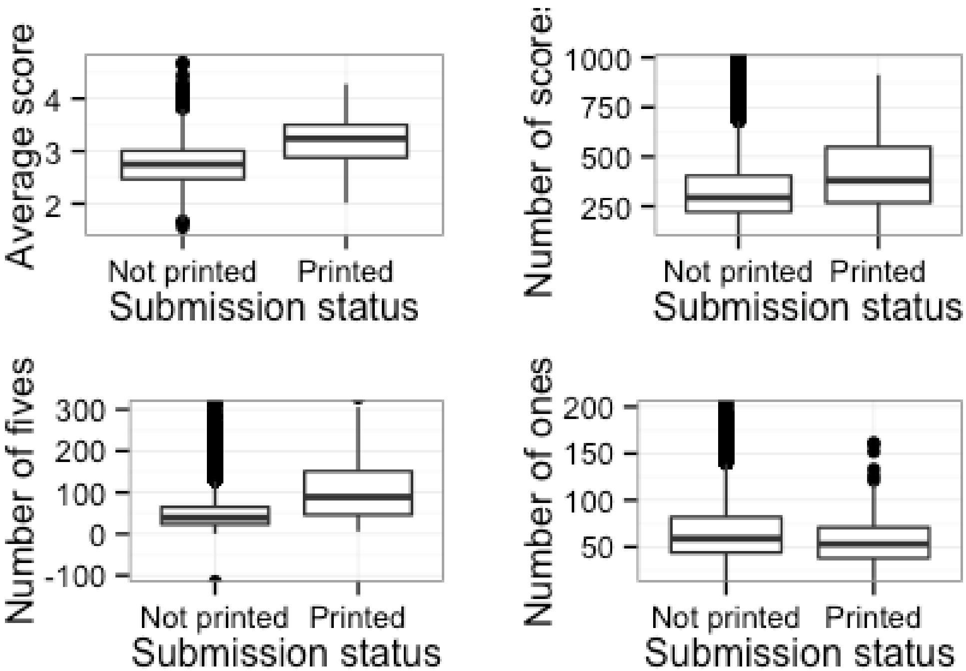


Distribution of fives for printed designs has fatter right tail than the distribution for designs that have not been printed. If design gathers more than 150 fives it appears to have good changes of getting printed.



Distributions for numbers of ones are almost identical for printed and not printed designs.

Boxplots



Based on the boxplots average score looks like the best predictor of designs getting printed. Number of fives is probably the second best, followed by number of scores. Number of ones appears useless in predicting the print status.

```
## Removed 2 observations.

##       avg.score score fives ones     printed
## 2919       2.25   124    13   44 Not printed
## 25684      3.03   243    48   38 Not printed
```

Logistic regression analysis

```
## glm(formula = printed ~ avg.score, family = binomial(), data = trai
n)
```

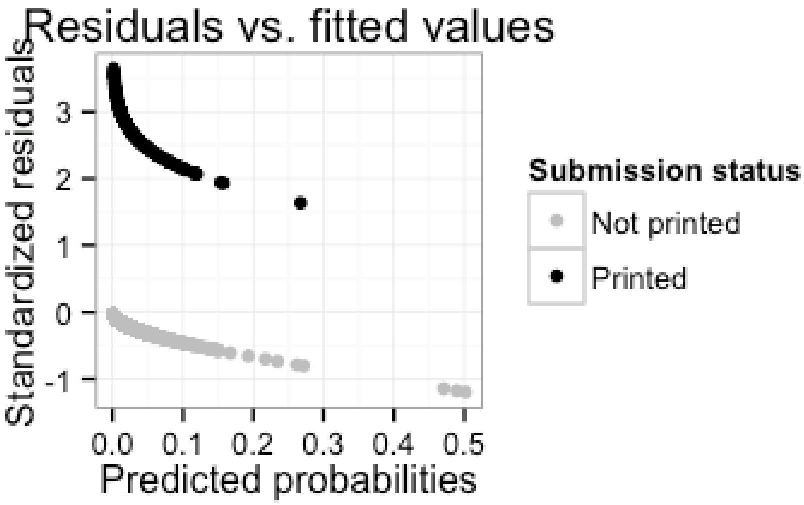
	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-11.62504	0.6530317	-17.80164	0
avg.score	2.48032	0.2063156	12.02197	0

The model estimates the probabilities of designs being printed based on the average score they have gathererd. It is statistically significant. The probability of getting this kind of data by change if there was no statistical effect between average score and design getting printed is practically zero.

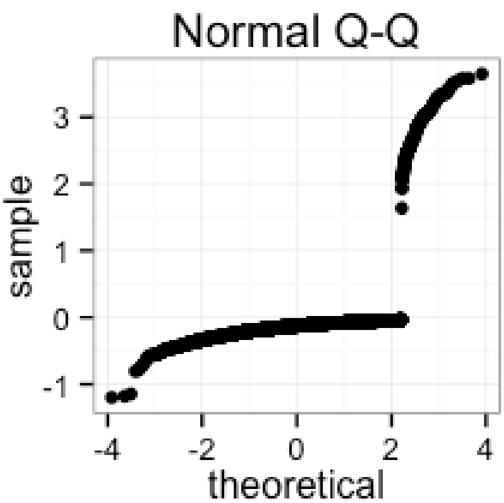
```
## modelChi       146.1841
## chidf          1
## chisq.prob     0
## Pseudo R^2 for logistic regression
## Hosmer and Lemeshow R^2    0.095
## Cox and Snell R^2          0.013
## Nagelkerke R^2             0.102
## Odds ratios:
## (Intercept)   avg.score
##       0.000      11.945
## Confidence intervals:

## Waiting for profiling to be done...

##             0.5 % 99.5 %
## (Intercept) 0.000  0.000
## avg.score   7.045 20.437
```

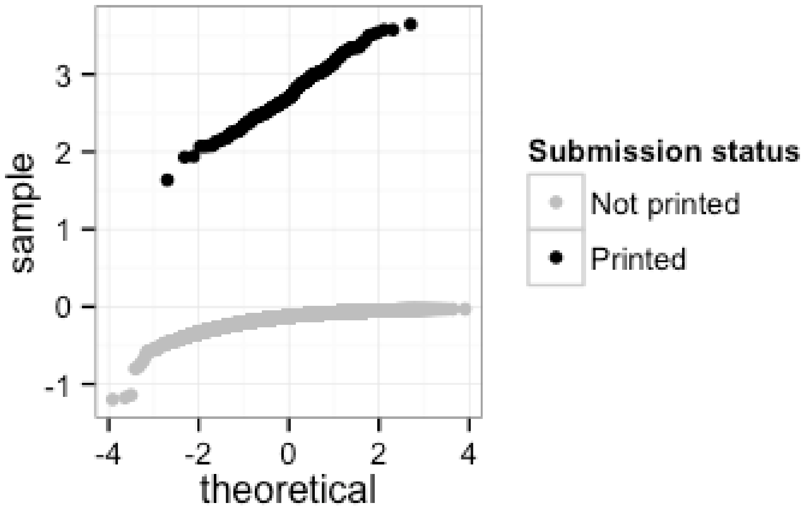
Chi-squared test shows the model fits the data significantly better than random change.Pseudo R^2 statistics still suggest that the effect is not very large. The average scores alone cannot explain which designs get printed. The odds ratio of average score is 15.7 with 99% confidence interval from 9.1 to 27.2. Getting 1 unit better average score increases the probability of design getting printed by 9 to 27 fold.

Model diagnostics



Designs that have not been printed have relatively low residuals. On the other hand printed designs have problematically large residuals. The model does not work well with the designs that do get printed.





Q-Q plot does not look good either, presumably because of the model's poor performance with printed designs.

```
## 141 observations have residuals larger than 2 standard deviations.
##  670 observations have leverage more than 2 times larger than the a
verage.
```

Observations with large residuals:

```
##
##         Not printed Printed
##   FALSE       10761       3
##   TRUE            0     141
```

All cases with problematically large residuals are printed designs.

Observations with large leverage:

```
##
##         Not printed Printed
##   FALSE       10142      93
##   TRUE          619      51
```

Observations with large leverage on the model are more equally distributed among printed and not printed designs.

```
## train$printed[train$large.leverage]: Not printed
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   3.390   3.440   3.500   3.554   3.620   4.690
## --------------------------------------------------------
## train$printed[train$large.leverage]: Printed
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   3.390   3.470   3.580   3.626   3.755   4.280

## train$printed[!train$large.leverage]: Not printed
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.650   2.450   2.710   2.700   2.968   3.380
## --------------------------------------------------------
## train$printed[!train$large.leverage]: Printed
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.010   2.710   2.940   2.901   3.210   3.370

sum(train$avg.score >= 3.39 & train$printed == "Not printed")

## [1] 619

sum(train$avg.score >= 3.39 & train$printed == "Printed")

## [1] 51
```

Both printed and not printed designs with high scores have large leverage. Perhaps this has something to do with the fact that most of the designs do not get printed.

Improvements to the model

```
##
## CORRELATIONS
## ============
## - correlation type:  pearson
## - correlations shown only when both variables are numeric
##
##           avg.score score fives   ones
## avg.score         . 0.361 0.519 -0.347
## score         0.361     . 0.717  0.609
## fives         0.519 0.717     .  0.184
## ones         -0.347 0.609 0.184      .
```

Correlations between variables do not look so large that multicollinearity would be a problem.

```
## glm(formula = printed ~ avg.score.norm, family = "binomial",
##     data = train)
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.7950678	0.1162445	-41.24985	0
avg.score.norm	0.9781646	0.0813647	12.02197	0

```
## glm(formula = printed ~ avg.score.norm + fives.norm, family = "bino
mial",
##     data = train)
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.7983911	0.1170067	-41.0095462	0.000000
avg.score.norm	1.0240016	0.1667004	6.1427663	0.000000
fives.norm	-0.0501329	0.1588824	-0.3155343	0.752356

```
## glm(formula = printed ~ avg.score.norm + fives.norm + score.norm,
##     family = "binomial", data = train)
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.8064824	0.1186115	-40.5229178	0.0000000
avg.score.norm	0.9310492	0.2354319	3.9546438	0.0000766
fives.norm	0.1377743	0.3735568	0.3688175	0.7122637
score.norm	-0.1227712	0.2213335	-0.5546889	0.5791075

```
## glm(formula = printed ~ avg.score.norm + fives.norm + score.norm +
##     ones.norm, family = "binomial", data = train)
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.8135734	0.1199468	-40.1309113	0.0000000
avg.score.norm	0.7226408	0.4507425	1.6032232	0.1088854
fives.norm	0.3318529	0.5191636	0.6392068	0.5226884
score.norm	-0.1040789	0.2247486	-0.4630902	0.6432997
ones.norm	-0.1298122	0.2392477	-0.5425849	0.5874156

Resid. Df	Resid. Dev	Df	Deviance	Pr(>Chi)
10903	1386.129	NA	NA	NA
10902	1386.029	1	0.0999022	0.7519470
10901	1385.722	1	0.3075847	0.5791661

10900	1385.425	1	0.2960694	0.5863571

After using average score to predict print status adding more variables does not improve the model. Neither number of fives or number of scores is statistically significant. Comparing the models using anova further confirms the lack of improvement.

```
## glm(formula = printed ~ avg.wo5.norm, family = "binomial", data = t
rain)
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.7340285	0.1151703	-41.10460	0
avg.wo5.norm	0.9419903	0.0906730	10.38887	0

```
## glm(formula = printed ~ avg.wo5.norm + fives.norm, family = "binomi
al",
##     data = train)
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.7781687	0.1172529	-40.750963	0.00e+00
avg.wo5.norm	0.5681336	0.1189873	4.774740	1.80e-06
fives.norm	0.4988207	0.1140422	4.374001	1.22e-05

Resid. Df	Resid. Dev	Df	Deviance	Pr(>Chi)
10903	1416.013	NA	NA	NA
10902	1399.143	1	16.86994	4e-05

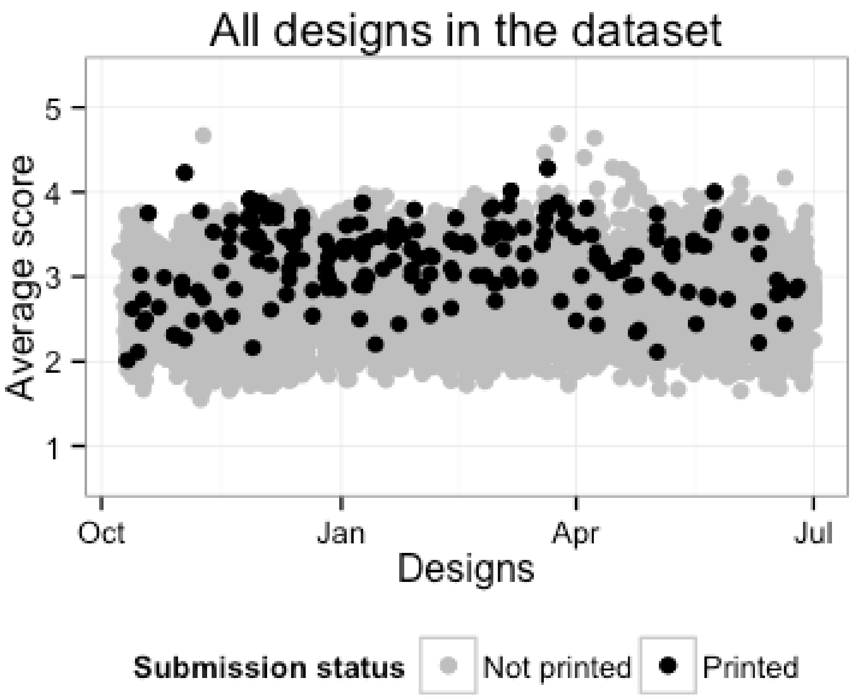
Here the logistic regression is first performed with average score with fives removed and then number of fives is added as a variable. This time the addition of number of fives improves the model.

```
## modelChi       146.1841
## chidf          1
## chisq.prob     0
## Pseudo R^2 for logistic regression
## Hosmer and Lemeshow R^2    0.095
## Cox and Snell R^2          0.013
## Nagelkerke R^2             0.102
## Odds ratios:
## (Intercept)   avg.score
##       0.000      11.945
## Confidence intervals:

## Waiting for profiling to be done...
```

```
##             0.5 % 99.5 %
## (Intercept) 0.000  0.000
## avg.score   7.045 20.437

## modelChi       133.1697
## chidf          2
## chisq.prob     0
## Pseudo R^2 for logistic regression
## Hosmer and Lemeshow R^2    0.087
## Cox and Snell R^2          0.012
## Nagelkerke R^2             0.093
## Odds ratios:
##  (Intercept) avg.wo5.norm   fives.norm
##        0.008        1.765        1.647
## Confidence intervals:

## Waiting for profiling to be done...

##              0.5 % 99.5 %
## (Intercept)  0.006  0.011
## avg.wo5.norm 1.319  2.431
## fives.norm   1.212  2.176
```

When compared to original model that used only average score, the original model has slightly higher pseudo R^2 values. Performance of the original model on the training set could not be improved by adding variables to the model.

As a conclusion, using average score to predict print status of designs provides the logistic regression model with the best fit.

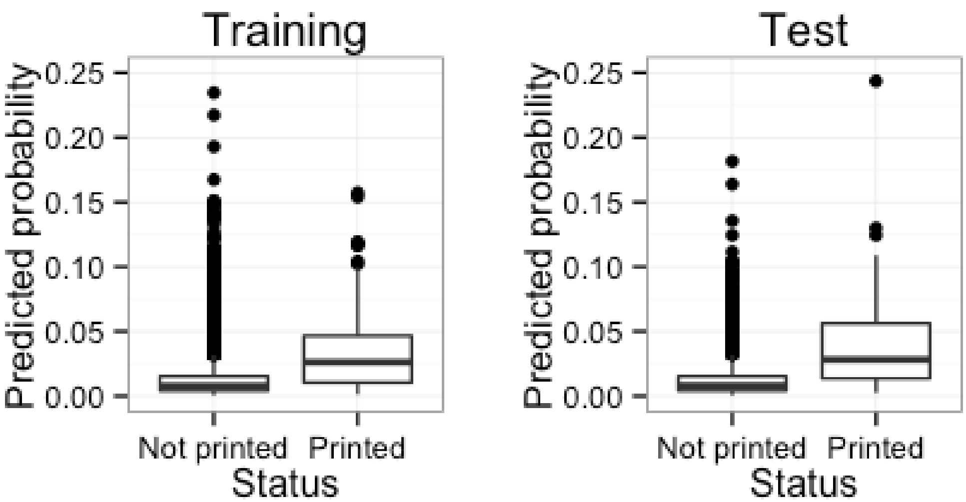


This graph shows that although there is a tendency for designs with higher average score to get printed more often, it is not possible to categorise them to printed and not printed designs based on the average score. The crowd can predict, but it cannot categorise.

Evaluation and validation

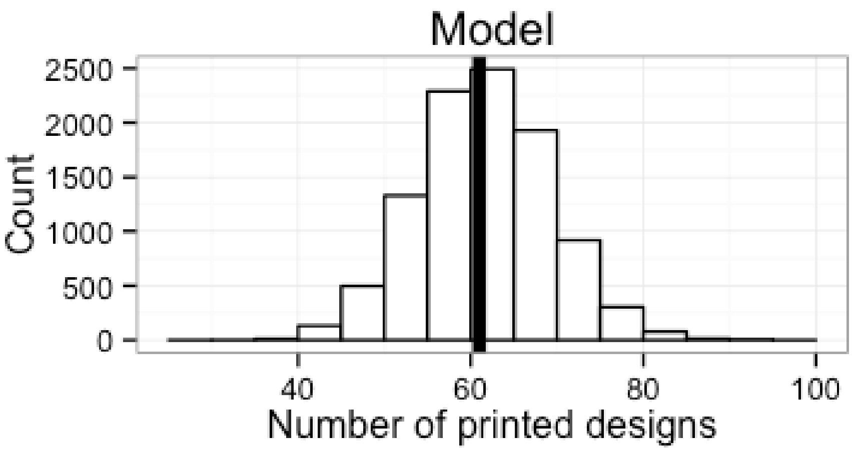
To validate the model the predicted probabilities between training and test sets are compared.

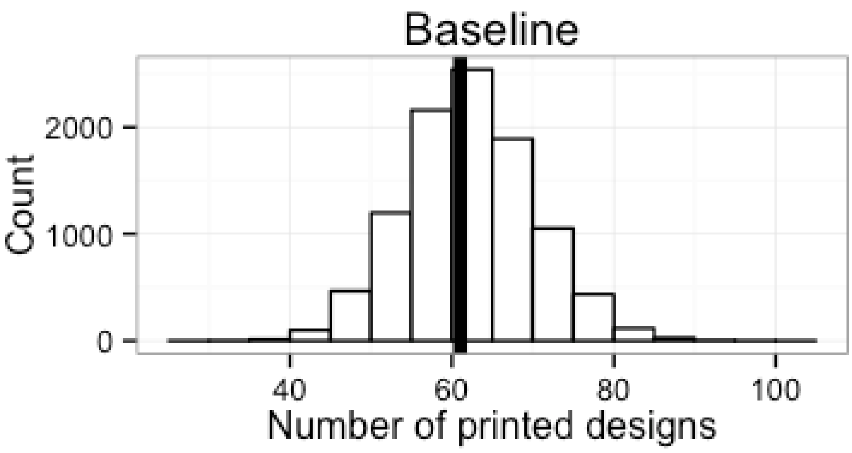
Mean squared error for the training set is 0.0127394 and MSE for the test set is 0.0123935. The difference is minimal, which indicates the model fits the training and tests sets equally well.



The printed designs in the test set tend to have higher predicted probabilities than designs that have not been printed.

Next a simple decision making simulation is used to further validate the model. It is assumed the decision makers choose designs to be printed according to probabilities predicted by the model. Decision making on the test set is simulated 10 000 times by assigning each design a random number drawn from a uniform distrtibution. If the number is smaller or equal to predicted probability, the design is printed. Number of printed designs and mean average score and standard deviation of printed designs is stored and compared to the actual observed values and results of baseline decision making simulation. In baseline simulation each design is given equal probability of getting printed based on the probability of randomly selected design being printed in the training set (number of printed designs in training set / number of design in training set).

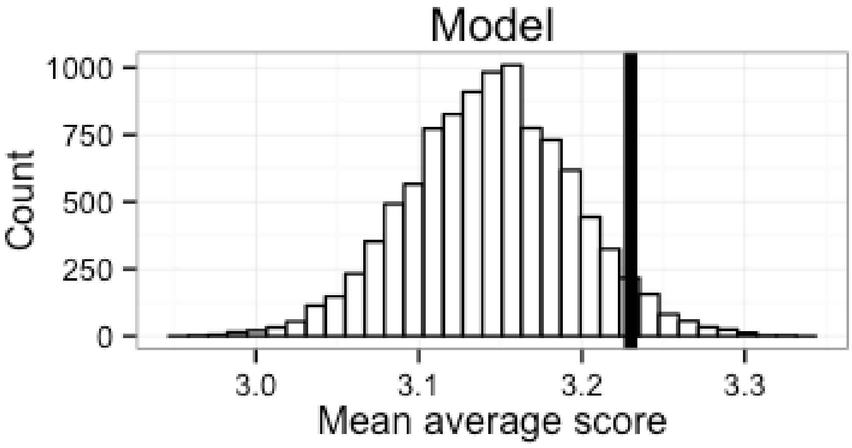


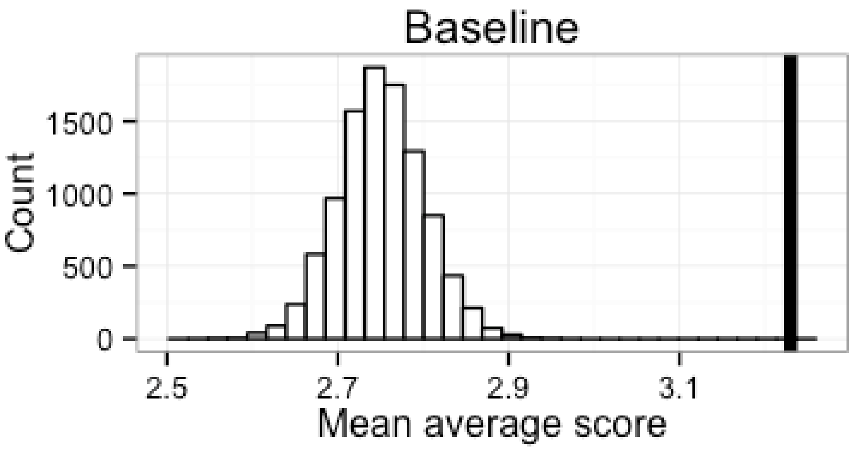


Both model and baseline simulation tend to produce the similar numbers of printed designs as actually observed, which indicates the ratio of designs chosen to be printed is similar in training and test sets.

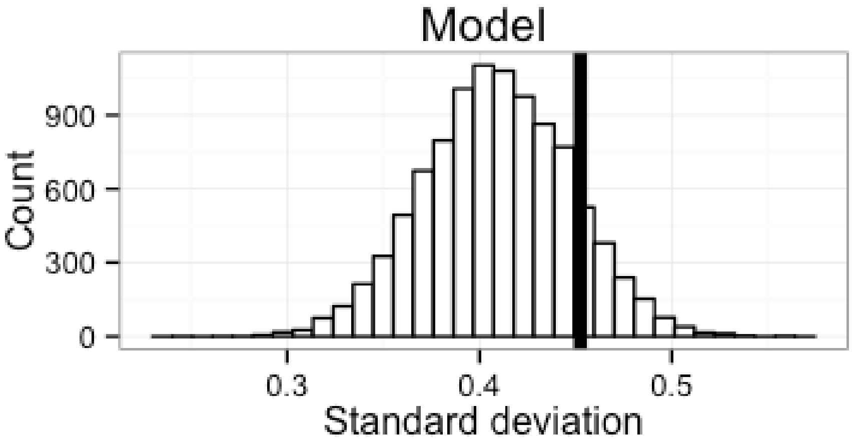
```
## Ratio of printed designs in training set is  0.0132
##  Ratio of printed designs in test set is  0.0131
```

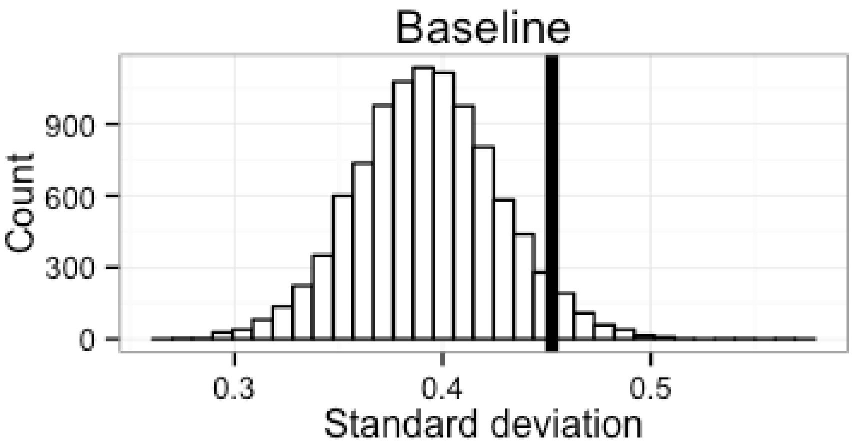
This is indeed the case.





Mean average score of printed designs in model simulation are centered around the actually observed mean average score. Baseline simulation never gets the correct value. Model is thus much better fit to the data than the baseline.





Regarding the standard deviation of average scores of printed designs the model simulation fares slightly better than the baseline simulation. It appears that in reality there is more variation in scores of printed designs than the decision making simulation typically generates.

Final model

Logistic regression model using average score to predict print status is fitted using the full dataset.

```
## glm(formula = printed ~ avg.score, family = binomial(), data = thre
adless)
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-12.129472	0.5592718	-21.68797	0
avg.score	2.640671	0.1756821	15.03096	0

```
## modelChi       232.4524
## chidf          1
## chisq.prob     0
## Pseudo R^2 for logistic regression
## Hosmer and Lemeshow R^2    0.106
## Cox and Snell R^2          0.015
## Nagelkerke R^2             0.113
## Odds ratios:
## (Intercept)   avg.score
##       0.000      14.023
## Confidence intervals:

## Waiting for profiling to be done...

##             0.5 % 99.5 %
## (Intercept) 0.000   0.00
## avg.score   8.949  22.15
```

634. KIVYIRO, PENDO. Foreign direct investment, clean development mechanism, and environmental management: a case of Sub-Saharan Africa. 2015. Diss.

635. SANKALA, ARTO. Modular double-cascade converter. 2015. Diss.

636. NIKOLAEVA, MARINA. Improving the fire retardancy of extruded/coextruded wood-plastic composites. 2015. Diss.

637. ABDEL WAHED, MAHMOUD. Geochemistry and water quality of Lake Qarun, Egypt. 2015. Diss.

638. PETROV, ILYA. Cost reduction of permanent magnet synchronous machines. 2015. Diss.

639. ZHANG, YUNFAN. Modification of photocatalyst with enhanced photocalytic activity for water treatment. 2015. Diss.

640. RATAVA, JUHO. Modelling cutting states in rough turning of 34CrNiMo6 steel. 2015. Diss.

641. MAYDANNIK, PHILIPP. Roll-to-roll atomic layer deposition process for flexible electronics applications. 2015. Diss.

642. SETH, FRANK. Empirical studies on software quality construction: Exploring human factors and organizational influences. 2015. Diss.

643. SMITH, AARON. New methods for controlling twin configurations and characterizing twin boundaries in 5M Ni-Mn-Ga for the development of applications. 2015. Diss.

644. NIKKU, MARKKU. Three-dimensional modeling of biomass fuel flow in a circulating fluidized bed furnace. 2015. Diss.

645. HENTTU, VILLE. Improving cost-efficiency and reducing environmental impacts of intermodal transportation with dry port concept – major rail transport corridor in Baltic Sea region. 2015. Diss.

646. HAN, BING. Influence of multi-phase phenomena on semibatch crystallization processes of aqueous solutions. 2015. Diss.

647. PTAK, PIOTR. Aircraft tracking and classification with VHF passive bistatic radar. 2015. Diss.

648. MAKKONEN, MARI. Cross-border transmission capacity development – Experiences from the Nordic electricity markets. 2015. Diss.

649. UUSITALO, ULLA-MAIJA. Show me your brain! Stories of interdisciplinary knowledge creation in practice. Experiences and observations from Aalto Design Factory, Finland. 2015. Diss.

650. ROOZBAHANI, HAMID. Novel control, haptic and calibration methods for teleoperated electrohydraulic servo systems. 2015. Diss.

651. SMIRNOVA, LIUDMILA. Electromagnetic and thermal design of a multilevel converter with high power density and reliability. 2015. Diss.

652. TALVITIE, JOONAS. Development of measurement systems in scientific research: Case study. 2015. Diss.

653. ZUBEDA, MUSSA. Variational ensemble kalman filtering in hydrology. 2015. Diss.

654. STEPANOV, ALEXANDER. Feasibility of industrial implementation of laser cutting into paper making machines. 2015. Diss.

655. SOKOLOV, MIKHAIL. Thick section laser beam welding of structural steels: methods for improving welding efficiency. 2015. Diss.

656. GORE, OLGA. Impacts of capacity remunerative mechanisms on cross-border trade. 2015. Diss.

657. AURINKO, HANNU. Risk assessment of modern landfill structures in Finland. 2015. Diss.

658. KAIJANEN, LAURA. Capillary electrophoresis: Applicability and method validation for biorefinery analytics. 2015. Diss.

659. KOLHINEN, JOHANNA. Yliopiston yrittäjämäisyyden sosiaalinen rakentuminen. Case: Aalto-yliopisto. 2015. Diss.

660. ANNALA, SALLA. Households' willingness to engage in demand response in the Finnish retail electricity market: an empirical study. 2015. Diss.

661. RIABCHENKO, EKATERINA. Generative part-based Gabor object detector. 2015. Diss.

662. ALKKIOMÄKI, VILLE. Role of service and data reuse in enterprises. 2015. Diss.

663. VÄNTSI, OLLI. Utilization of recycled mineral wool as filler in wood plastic composites. 2015. Diss.

664. KLEMOLA, KATJA. Tuottavuuden, vaikuttavuuden ja kustannusvaikuttavuuden arviointi alueellisesti integroiduissa sosiaali- ja terveyspalveluissa – palvelujen käyttöön perustuva malli ja esimerkkejä. 2015. Diss.

665. HEROLD, KRISTIINA. Impact of Word-of-Mouth on consumer decision-making: An information processing perspective in the context of a high-involvement service. 2015. Diss.

666. OLABODE, MUYIWA. Weldability of high strength aluminium alloys. 2015. Diss.

667. VANHALA, ERNO. The role of business model in computer game development organizations. 2015. Diss.

668. SALAMPASIS, DIMITRIOS. Trust-embedded open innovation: Towards a human-centric approach in the financial industry. 2015. Diss.

669. DE SMET, DIETER. Innovation ecosystem perspectives on financial services innovation. 2015. Diss.

670. PORRAS, PÄIVI. Utilising student profiles in mathematics course arrangements. 2015. Diss.