HOME ABOUT LOGIN REGISTER SEARCH CURRENT
ARCHIVES ANNOUNCEMENTS SUBMISSIONS

Home > Volume 24, Number 10 - 7 October 2019 > Eklund

The crowd in crowdsourcing: Crowds as a pragmatic research method

by Lina Eklund, Isabell Stamm, and Wanda Katja Liebermann

Abstract

Crowdsourcing, as a digital process employed to obtain information, ideas, and solicit contributions of work, creativity, etc., from large online crowds stems from business, yet is increasingly used in research. Engaging with previous literature and a symposium on academic crowdsourcing this study explores the underlying assumptions about crowdsourcing as a potential academic research method and how these affect the knowledge produced. Results identify crowdsourcing research as research *about* and *with* the crowd, explore how tasks can be *productive*, *reconfiguring*, and *evaluating*, and how these are linked to *intrinsic* and *extrinsic* rewards, we also identify three types of platforms: *commercial platforms*, *research-specific platforms*, and *project specific platforms*. Finally, the study suggests that crowdsourcing is a digital method that could be considered a *pragmatic* method; the challenge of a sound crowdsourcing project is to think about the researcher's relationship to the crowd, the tasks, and the platform used.

Contents

[Introduction](#)
[Approach and material](#)
[Methods and scientific knowledge production](#)
[Harnessing the crowd to produce knowledge](#)
[Methodological implications of the crowdsourcing process](#)
[Concluding discussion: Crowdsourcing as a pragmatic method](#)

Introduction

So, start to think about what we can do — *when*, not if — and ways in which you can mobilize crowds to do something for the public good or improve your research along the way.
 — Daren Brabham, 6 November 2015.

A large group of researchers gathered for a conference on the University of California, Berkeley campus are listening to a presentation and waiting for

[OPEN JOURNAL SYSTEMS](#)

[Journal Help](#)

USER

Username

Password

☐ Remember me

JOURNAL CONTENT

Search

Search Scope

All

Browse

- [By Issue](#)
- [By Author](#)
- [By Title](#)
- [Other Journals](#)

FONT SIZE

CURRENT ISSUE

ATOM	1.0
RSS	2.0
RSS	1.0

ARTICLE TOOLS

[Abstract](#)

[Print this article](#)

[article](#)

[Indexing](#)

[metadata](#)

[How to cite item](#)

[Email this](#)

article (Login required)

[Email the](#)

author (Login required)

ABOUT THE AUTHORS

Lina Eklund
<http://www.sirg.se>

Uppsala University
 Sweden

their turn to view a children's book — entitled *The Adventures of Hashtag the Snail* (<https://ioanaliterat.com/hashtag/>) — that is being passed around. The book tells the story of a little snail called Hashtag, who, on his way home, stumbles upon an empty snail shell. This is the narrative launch for a literary project using a novel technique for developing a story about the shell's former inhabitant. The group's fascination with Hashtag has less to do with the charming story and illustrations than with the way the book was created. The presenter and project author Ioana Literat describes it as a children's book 'about the Internet by the Internet.' The book is the product of the collective intellectual labor of anonymous participants using an online crowdsourcing platform. This colorful printed object offers tangible proof of the capacity of a virtually configured group to join together to produce a creative artefact — a work of the crowd.

The enthusiasm over this book reflects the hopes that researchers have for the potential of crowdsourcing applications in research; if the crowd can create something like a book, the potential for research application seems vast. Indeed, as a sociotechnical practice it seems to excite new possibilities for and challenges to scientific knowledge production beyond the scope and scale of traditional research projects (Estellés-Arolas and González-Ladrón-de-Guevara, 2012; Pedersen, *et al.*, 2013; Tarrell, *et al.*, 2013; Malone, *et al.*, 2009; Geiger, *et al.*, 2011).

Crowdsourcing is a digital process employed to obtain information, ideas, and solicit contributions of work, creativity, and so forth, from large online crowds (not to be confused with crowdfunding, an online monetary funding technique). Crowdsourcing has been going on for years on sites such as threadless.com (<https://www.threadless.com>) where people design t-shirt prints and then vote to decide which designs will go to market. Often crowdsourcing is thought of as a method to deploy human power to work with big data in cases computers cannot, but as the #hashtag project shows, this is only one possibility. In essence, crowdsourcing harnesses the time, energy, and talents of individuals, whom we call crowd-taskers, reached through the Internet to perform a pre-given undertaking (Shepherd, 2012). We use the term "crowd-tasker" to open up the meaning of participation beyond what terms such as crowdworker, or, from the realm of academic research — subject, interviewee, participant — have traditionally connoted.

In some ways it is the digital version of citizen science. Its hybrid etymology — "crowd" plus "outsourcing" — signals ambiguous potentials: from the exacerbation of dispersed, part-time, tedious, piecemeal-based labour conditions enabled through Web platforms (Irani, 2015) to, more optimistically, new digitally-aided structures for conducting large, complex, collaborative, and interactive projects (Brabham, 2013). Borrowing from businesses that use crowds to design, innovate, and produce, over the last decade, researchers in the humanities and social sciences have begun to reconfigure the landscape of academic knowledge production. Researchers now engage crowds to code large data sets, participate in online experiments, create children's books, translate ancient texts, and much more. In this study we understand crowdsourcing as a digital procedure by which researchers engage with large numbers of individuals reached through an Internet platform.

The novelty of this technique of research has yet to produce routines and best practices, and, more importantly, consensus is lacking — particularly, between disciplines — about what this "method" is, how to use it, and even why to use it. The majority of existing literature on crowdsourcing in academic research is narrowly focused on technical, procedural, and efficacy questions, such as quality control measures (Hansen, *et al.*, 2013; Daniel, *et al.*, 2018), and recruitment and retention of participants (Prestopnik and Crowston, 2013; Robson, *et al.*, 2013). To date, there is scant critical scholarship on crowdsourcing as a research method, considering its epistemological implications (Pedersen, *et al.*, 2013), or in other words, probing the nature of the knowledge produced with this approach. We would expect crowdsourcing as a scientific research method to differ from its uses in commercial, civic, and other sectors, because scholarly knowledge has more defined criteria for establishing valid knowledge.

Crowdsourcing is not native to research and its origin in business comes with particular and often hidden assumptions about the world, which directly influence the knowledge produced. We thus ask, does crowdsourcing qualify as a scientific method and how can we understand its methodological underpinnings? In this study, we aim to demystify crowdsourcing as it is used in humanities and social science research and connect that discussion to a broader debate about scientific knowledge

Assistant Professor
at the Department
for Informatics and
Media.

Isabell Stamm
Freigeist-Research
Group, Technical
University in Berlin
Germany

Wanda Katja
Liebermann
School of
Architecture, Florida
Atlantic University
United States

production. This link is significant, as the business origins of the method potentially shape the knowledge generated by the method (Marres and Weltevrede, 2013). Drawing from digital methods discourse and philosophy of science, we discuss important methodological implications of key ideas circulating about crowdsourcing. We argue that as researchers and academic knowledge producers, we should not forget the parameters of knowledge production. We need to think about and reflect on the methodological underpinnings of new digital methods.

Approach and material

The starting point of our study was a symposium the three authors organized on scientific crowdsourcing held at University of California, Berkeley in 2015, "Crowdsourcing and the Academy: Exploring promises and problems of collective intelligence methods in humanities and social science research" (<https://hssa.berkeley.edu/crowdsourcing-symposium>). The symposium was the point of departure for this investigation. Since our interest lay in how academics engage with, think about, and discuss crowdsourcing as a method the symposium offered us rich and relevant data. A central aim of the symposium was to theorize the connection between project design and knowledge production in crowdsourcing, which was reflected in the two panels, which followed an introductory keynote by Daren Brabham of the University of Southern California Annenberg School for Communication and Journalism, who defined crowdsourcing and laid out the current landscape. The panels: "Talking from Experience," featured scholars discussing their work using crowdsourcing methods; "Theoretical Considerations," explored the cultural and social implications of crowdsourcing in the humanities and social sciences.

The first panel included Ioana Literat's dissertation project, mentioned earlier, a theorized process of creating a children's book, written and illustrated entirely by Amazon Mechanical Turk participants. The depth of her analysis and the volume of input — she collected 4,200 examples of written and visual material — highlights the breadth of possibilities for Internet-facilitated creativity. Next was the Perseids Project (<https://www.perseids.org>), an online platform for collaborative editing, annotation, and publication of digital texts and scholarly annotations. This project, presented by Tim Buckingham of Tufts University and Simona Stoyanova of the Open Philology Project of the University of Leipzig, combines contributions from the classroom and the general public, grants contributor authorship, and uses a complex review and feedback system. Third, the Deciding Force Project (<https://www.decidingforce.org>), led by Nick Adams of the Berkeley Institute for Data Science (BIDS), produced its own software called Text Thresher (<https://bids.berkeley.edu/research/text-thresher>), which combines crowdsourcing, machine learning, and content analysis in order to conduct a 'comparative study of protest policing.' Using the U.S. Occupy Movement, this project collects, classifies, and analyses an enormous dataset of newspaper articles chronicling police and Occupy Movement interaction. Finally, UC Berkeley School of Information's Marti Hearst offered insights into how to facilitate peer evaluation and co-learning among crowd workers in order to enable more challenging tasks in her presentation 'Improving crowdwork with small group discussions'. These projects do not represent the full breadth of crowdsourcing in research but, rather, each illuminate significant aspects of how crowdsourcing challenges traditional norms of and opens new possibilities for academic research.

In the second panel, discussants reflected critically on the challenges and promises of crowdsourcing as a method, drawing from their own experience. Lily Irani of the University of California, San Diego questioned the constitution of "the crowd," drawing attention to the limited insights we have about the demographics and working conditions of individuals performing crowdsourced tasks. Along these lines, Trebor Scholz of the New School for Liberal Arts, elaborated on the broader cultural implications of sharing practices based on digital technologies. Finally, Stuart Geiger of UC Berkeley's School of Information pointed to parallels between crowdsourcing and citizen science, in the sense that creating a knowledge community raises questions of authorship and copyright for traditional scientific knowledge production.

We had access to transcriptions of the eight talks including the keynote speech, panel presentations, and panel and audience discussion, which were recorded with the permission of the organizers and presenters. Our analytical approach was inspired by grounded theory, using topic

categorization to find common themes and trends in source material. The interview was coded manually and inductively, with codes and arguments rising from the data. We identified three critical areas in this emerging research practice: 1) crowd-taskers as both subject and objects; 2) research tasks; and, 3) the influence of platforms. We engaged each of the emerging themes with relevant current topics and arguments in a social science discourse about methodology. We did many rounds of writing, thinking, and talking through our data, keeping in constant dialogue with both our data and the discourse on knowledge production. What emerged is a heuristic that allows us to reflect upon crowdsourcing as a method, uncovering its methodological implications. We identified one additional key area, namely the ethical implications of this method, but we chose to exclude ethics from this article, as the topic was broad enough to warrant its own separate investigation. We do, however, want to highlight that ethics of crowdsourcing is a topic that researchers need to engage with in the future.

Methods and scientific knowledge production

When setting out to study the empirical world, a researcher ideally first creates a research design, which is commonly thought of as a 'plan of action that links philosophical assumptions to specific methods' [1]. According to Blumer (1969), a research design includes developing a picture of the empirical world, asking questions about that world and turning these into researchable problems. The next step is finding the best means of doing so, including making choices about methods, the development and use of concepts, and the interpretation of findings (Blumer, 1969). In practice, many researchers have methods they prefer, which are often determined before research questions are designed. A research method is a set of techniques or heuristics that defines a system for studying a phenomenon. These techniques and heuristics guide researchers in the creation of knowledge and they vary depending on the traditions of the specific subject field (Alasuutari, *et al.*, 2008). With methods we can create a research design. Of significance is that all methods contain underlying assumptions about how we can come to understand the empirical world, and the understanding that different methods will result in different types of scientific knowledge; in other words, our epistemological understanding shapes outcomes.

When philosophy of science attempts to answer questions about methods and the nature of scientific knowledge, two key research paradigms have dominated in the social sciences: positivism (realism) and constructivism (relativism). Paradigms here represent established standards, a certain way of viewing the world (Burrell and Morgan, 1979). Positivists tend to prefer scientific quantitative methods with large scale surveys in order to get an overview of society and study trends, 'laws', and structures which guide human behaviour. Positivism sees ideal research and scientific knowledge as deductive, objective, and value-free. Constructivism rose as a critique of positivism and was influenced by the idea that knowledge cannot be separated from the knower (Steedman, 1991). Constructivism has moved research from ideas of how to create true, objective knowledge to a more reflexive approach. It favors humanistic qualitative methods and argues that people experience and understand the same 'objective reality' in very different ways and have their own reasons for acting in the world. Constructivism views research and knowledge as fragmented, inductive, and subjective. In the humanities, interpretative methods that lean more towards the qualitative side of the methodological spectrum have long dominated. However, digital humanities is a new frontier, making thus-far-unimaginably enormous data available to research. Particularly, machine learning promises a way to reconstruct meaning even for large volumes of text (Wiedemann, 2013). This digital turn has triggered a rise of computational and mathematical methods among humanists as well as new methods in the social sciences.

Each paradigm's view of research participants diverges also. Quantitative research assumes that research participants have the same view of the questions asked as the researcher does. Even qualitative research often operates under the assumption that agreeing to take part in a study means agreement with the boundaries defined by the researcher. Surveys frequently contain an implicit assumption that respondents will and do understand and interpret the questions posed in the same way as the researchers do, that they have attitudes about the questions asked and that they want to share these (Feilzer, 2010). Constructivist research,

however, often pays attention to what is socially constructed and thus how scientific knowledge will be affected by the chosen method and the relationship it imposes between researcher and research subjects.

The paradigms of positivism and constructivism have, at times, been at odds as they propose two different ways of looking at what constitutes scientific knowledge: positivists search for the one truth and constructivist study how the world is perceived and thus the many truths that exist. But some researchers have come to find this division limiting and have argued that each paradigm, while promoting very different types of research, has its own capacities to ask, and answer, particular research questions. Moreover, Dewey (1925), a foundational pragmatist theorizer, even argued that the main research paradigms of positivism and subjectivism come from the same paradigm family, as they both search for the truth, "whether it is an objective truth or the relative truth of multiple realities" [2]. In response, pragmatism has gained traction as a research philosophy that could put an end to the old dualism of qualitative and quantitative or data-driven versus interpretative methods (Baert, 2005; Dewey, 2007). Pragmatism as a research philosophy aims "to set aside considerations about what is ultimately true in favour of what is ultimately useful" [3]. Pragmatism considers things relevant only if it supports action, and acknowledges that there are multiple ways of interpreting the world and doing research. Pragmatism further acknowledges how a method of inquiry may alter the scientific knowledge produced; in this sense, knowledge is seen as active. In the pragmatist tradition, the research question is the most important factor as it will determine a scientist's research plan and thus methodological approach. Method then, in a pragmatic sense, always entails the question of aims, and therefore partially depends on what a researcher wants to achieve (Baert, 2005). Pragmatism has been particularly influential in the mixed methods debate because it has the potential to bridge between different research traditions (Feilzer, 2010). Using a pragmatic approach, we can see methods as active agents in shaping knowledge, without having to assume an underlying 'correct' method. Pragmatism instead frees the researcher from the artificial choice among the two paradigms [4].

In sum, various paradigms for understanding scientific knowledge see different "correct" ways of doing research. Every method is embedded in methodological considerations, and whether or not a method is considered as adequate or even legitimate depends upon the underlying paradigm. Hence, applying crowdsourcing in research requires reflection on the characteristics that make crowdsourcing a method as well as on foundational understandings of knowledge production.

Harnessing the crowd to produce knowledge

Howe's original definition of crowdsourcing, 'the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call' (Howe, 2006), which appeared a decade ago, in *Wired*, reminds us that crowdsourcing is a tool originally developed for commercial purposes, which launched businesses like iStockphoto.com and threadless.com. Crowdsourcing as a tool by which business organizations can harness the labor of anonymous and unattached task-contractors instead of by an employed and protected workforce, increasingly gains research attention. In parallel, studies that examine the socio-demographic composition of the crowd and the working conditions of the individual crowd-taskers have proliferated (Brabham, 2012; Kaufmann, *et al.*, 2011; Ross, *et al.*, 2010).

We differentiate this research about the crowd from research *with* the crowd, *i.e.*, study designs that actually apply the tool originally developed in the business realm for research purposes. The sociotechnical practice of crowdsourcing offers, at least on paper, the opportunity to access, generate, and analyze new kinds of data, at new scales, and in new ways. Crowdsourcing can be used to conduct data, for example by recruiting participants in survey studies or by having crowd-taskers collect pictures and snippets; crowdsourcing can be used to qualify data, for example write captions, or research specified information; crowdsourcing can be used to analyze data, for example, crowd-taskers can code text passages or extract information. In other words, crowdsourcing offers a procedure that guides researchers in the creation of knowledge. Crowdsourcing thus generally qualifies as research method (Snee, *et al.*, 2016). More specifically, crowdsourcing is a digital method in that it uses online and

digital technologies to collect and analyse research data (Snee, *et al.*, 2016). Following Rogers (2015; 2013), crowdsourcing, like Web crawling, Web scraping, and folksonomy, can be classified as an indigenous digital method as it is a Web technique deeply embedded in the online environment. This contrasts with analog methods that have become digitized, such as online surveys or online ethnography, traditional methods which have migrated online.

The peculiarity of crowdsourcing is its malleability. Research with the crowd can be implemented at different phases in the research process (data conduction, data analysis), with the crowd appearing as object of research, *i.e.*, as a specific group tapped to inform our understanding of the social world, or as subject in research, *i.e.*, as capable individuals assisting in the process of knowledge production. During the symposium it was acknowledged that crowdsourcing has the potential to be both a qualitative and quantitative method, as the following mention of sentiment analysis illustrates. The quote also argues that researchers in crowdsourcing often find it hard to acknowledge that the method is used in various epistemological contexts, something which could be a resource:

'Stuart: Think about it being objective versus being subjective. [...] I think we need to start to take that into account and be thinking about what kind of benefits we can bring by seeing different people who have different ideas about how to evaluate sentiment. Leveraging that as a resource as opposed to seeing that as, "oh my god this shuts down my way of thinking about it".'

At the same time, as in the quote, the symposium made obvious that researchers struggled with talking about crowdsourcing in the way we normally talk about research methods. In particular, conversation about what kind of knowledge crowdsourcing can produce or what types of questions it is suited for was limited. Marres and Weltevrede (2013) caution us about the translation of crowdsourcing into research contexts, observing that its application as a method requires critical reflection on the built-in values inherited from a commercial process and thus to question the data it generates. For digitized research methods such as online surveys or online ethnography, researchers have successfully revised their underlying assumptions of knowledge production [5]. Yet, we have just begun to probe the philosophical framework underlying the foundations of crowdsourcing as a method (Rogers, 2013). As crowdsourcing continues to permeate academic research it becomes increasingly important to recognize that, as with any methodology, crowdsourcing carries specific assumptions of the knowledge produced. During the symposium discussions circled around three interrelated topics, the composition of the crowd, the tasks, and platforms. In the following we will reflect upon the knowledge crowdsourcing can produce considering each of these topics.

Methodological implications of the crowdsourcing process

Crowd-taskers and the crowd

The open call nature of crowdsourcing leaves control over who participates mostly outside the researcher's control, which inherently leads to uncertainty about the composition of the crowd. This fact has implications for research that differs from civic, commercial, and artistic applications.

Studies *about* the crowd have been trying to disperse some of the mysteries about this group. The trope that anyone can be a crowd-tasker is prominent in crowdsourcing rhetoric: individuals can perform tasks on the go, anywhere, anytime, as long as they have access to digital technologies. However, the digital divide still exists. Many areas of the world have inadequate or no access to the Internet; and even in countries with almost universal access, some individuals will be more Internet savvy than others. Crowdsourcing platforms also play on the image of a vast, diverse, and global crowd in their marketing (Eklund, *et al.*, 2017). In reality studies on the identity of crowd workers and their working conditions have shown that the crowd is far less diverse than often assumed (Irani and Silberman, 2013; Willett, *et al.*, 2012; Literat, 2012; Zittrain, 2008). In early business applications it was widely assumed that the crowd consisted of amateurs, people with no skill or training in the specific task, who participated because it seemed fun or because they were motivated by monetary rewards. In fact, research on Amazon Mechanical

Turk shows that “Turkers” are more highly educated than the average American (Ross, *et al.*, 2010). The myth of the amateur crowd is often just that, a myth (Brabham, 2012). At the same time, there is also the “Wikipedia problem,” an environment where the majority of contributors are tech savvy white men (Glott, *et al.*, 2010).

A central goal of the Perseids Project is to translate ancient inscriptions into as many languages as possible. Therefore, the project was designed to reach skilled crowd-taskers from diverse cultural backgrounds. By contrast, the Next Stop Design Project, an effort to create the Salt Lake City, Utah bus stop shelter designs (<http://nextstopdesign.com>), was designed with a local Utah audience in mind, open to anyone willing to participate. While the Perseids Project succeeded in attracting the intended crowd, the bus stop project was in for a surprise. Brabham explained:

“We figured we would get just people from Utah, or maybe people who liked to ski and visited Utah a lot and had something to say about the transit hub, but, of course, it is the Internet. And an architectural competition blog in Germany picked it up, the blog of Google SketchUp, their official blog, picked it up, and all of a sudden, we had the whole world, which we did not intend; we did not even want, frankly. ‘Cos the point was to get Utah’s voice on this.”

Thus, a series of unpredicted digital events resulted in design contributions from all over the world, greatly improving the quality of submissions — but at a cost. Both projects revealed that crowd-taskers possess a wide array of specialized training and expertise, such as architecture or Latin and Greek language proficiency. Yet, who will accept a task, as crowdsourcing project are communicated online, is unpredictable. We therefore argue that determining a universal definition of the crowd is not possible as each project design instantiates a unique composition of crowd-taskers, something that cannot be entirely predicted beforehand.

This circumstance has implications for doing research with the crowd. The uncertainty about who will participate allows for a great deal of uncertainty about the nature of the crowd; which creates unique difficulties for this method. Moreover, the fact that it might produce its own demographic, as in Wikipedia (Glott, *et al.*, 2010) hinders both representativeness and democratic values. In other types of research, researchers exert a measure of control of who is included in studies, using controlled random sampling or by meeting informants face-to-face. However, the anonymous nature of the Internet eliminates these controls. Indeed, regaining control over crowd composition is a central promise of platforms that offer crowdsourcing services to researchers, a question to which we will return below. In actuality, we run the risk that myths about the crowd create proxy meanings that affect project design.

Crowdsourcing promises to scale up research to previously unreachable magnitude by its access to large crowds. However, in studies striving to uphold scientific rules of representativeness that conceptualize the crowd as any type of representative sample is problematic precisely due to our lack of control over who the crowd is. Interestingly, research on crowdsourcing has found that it leads to greater response diversity for survey studies than traditional student participant pools (Behrend, *et al.*, 2011). Nevertheless, the loss of control over who participates in a crowdsourced research project poses problems for studies based on experimental project design and other approaches that demand random sampling. Moreover, there is the risk of crowdsourcing creating its own demographic where some people will not be part of the crowd. For example, Wikipedia has a serious problem in attracting contributions from women (Menking and Erickson, 2015). The unpredictability and homogeneity of the crowd challenge scientific rules of representativeness, as well as ideals about democratic participation, thereby ruling out or greatly restricting the appropriateness of crowdsourcing for some research questions.

For interpretative research, in particular, an element of unpredictability can often be beneficial, especially in studying less explored knowledge terrains (Kitzinger, 1995). An empowered crowd can produce unanticipated and fortuitous research results. In this perspective, crowd-taskers turns into co-creators, or as Stuart Geiger suggested at the symposium, into ‘research assistants’. Arguably, the latter practice of crowdsourcing can be read as a digitized version of citizen science, which promotes research collaborations between scientists and volunteers, particularly (but not

exclusively) to expand opportunities for scientific data collection and to provide access to scientific information for members of the public. However, crowdsourcing cannot be limited to this one version. Moreover, the distance inherent to crowdsourcing — researchers do not meet research participants, who are almost always separated by digital interfaces — also promises to pose difficult questions for interpretative research to untangle. The differing underlying assumptions about who the crowd is and the flattening of research participants into a “crowd” go along with epistemological assumptions about the relationship between knower and the known, that is, power relations in knowledge production. Drawing on Wexler’s (2011) studies we can see that when crowdsourcing ignores the disruptive potential of crowds, it glosses over potential interpretations, arguments, and problems of the method.

Indeed, the great draw with crowdsourcing is the method’s ability to draw on large numbers of individuals. Because of the complexities of managing huge numbers of people, crowdsourcing reduces them to a faceless crowd. Instead of having to deal with each individual member, a researcher’s interaction is with the crowd itself; this is the essence of what crowdsourcing allows. In industry, such reduction of complexity is seen as a benefit, but when applied to research, it becomes inherently problematic, as it contradicts the basic idea that investigators control who participates in studies, either as part of the sample or as part of the project team.

Research tasks and assumed capability of the crowd

The task that researchers entrust a crowd to perform can take many forms and relates directly to assumptions about the capacity of the crowd. The breadth of symposium projects illustrates this perfectly: designing a local bus stop (Next Stop Design Project); annotating ancient texts from high-definition scanned images of objects, such as stone tablets and vases (Perseids Project); illustrations for a childrens book (*The Adventures of Hashtag the Snail*); and, coding text from newspaper articles about the Occupy Movement (Deciding Force Project).

Drawing on the symposium and crowdsourcing literature, we identify three broad categories of cognitive efforts that crowdsourcing projects ask of crowd-taskers. First, *productive tasks* involve the generation of ideas, designs, data, or text, including a wide range of crowd produced material such as pictures, drawings, sentences, stories, or completed surveys falls into this category — raw material that complements both constructivist and positivist research traditions. Second, *reconfiguring tasks* require the translation of original material into higher order concepts using a predefined interpretation scheme. Here, the researcher provides original material and asks the crowd to describe, tag, locate, annotate, code, or interpret the given material. Carletti, *et al.* (2013) make a similar distinction between these two types of crowdsourcing tasks in the digital humanities. A third category is *evaluating tasks*, which ask the crowd to assess the output of previous productive and reconfiguring tasks, creating a feedback loop.

Crowdsourcing assumes that there is potential to harness a dispersed collective intelligence, which under the right conditions, can be applied to solve problems (Wexler, 2011; Surowiecki, 2005). Yet intelligence implicitly and explicitly attributed to the crowd varies dramatically by the parameters of a given crowdsourcing project. At the symposium, Marti Hearst voiced a general consensus among researchers that a task can easily become too intricate or large to make sense for crowdsourcing. At what point a task becomes too complex, however, is subject to the individual researcher’s judgment and his or her assumptions about the crowd.

Increased task complexity goes along with a decrease in the knowledge authority of the researcher and necessitates heightened trust in the crowd’s capabilities. For example, the task of coding text passages in news articles requires human cognitive skills that surpass the current capabilities of computational coding and machine learning. As Nick Adams said ‘Humans are great at understanding meaning.’ Leaving the coding of newspaper articles to an untrained crowd, however, poses many risks, according to the researcher. Individual differences in coding approaches make standardization difficult and the quality of coding will likely vary. These problems lessen the researcher’s control over results and undermine standards of good research. For his Deciding Force Project, Adams approached these risks with two main simplifying strategies. First, as other authors have noted as a common trait to crowdsourcing projects (*e.g.*, Estellés-Arolas and Gonz´lez-Ladrón-de-Guevara, 2012), Adams broke tasks into smaller steps, *i.e.*, coding sentences or phrases instead of whole articles. Second, he used a reading comprehension task format for the

coding process. As he explained: 'I do not have to train people to do this task I just have to say, "hey, remember that reading comprehension protocol that you ran a thousand times through grade school" — I want you to do that' (Adams, 2015). In this way, Adams increases the comparability and standardization of results, and pulls the locus of control back to the researcher.

Researchers who try to develop techniques for enabling crowd creativity through more complex tasks (Kittur, *et al.*, 2011) view such simplification strategies as a reflecting distrust of the crowd's abilities. Simple and narrowly defined tasks tend to assume little creativity and cognitive capacity (Estellés-Arolas and González-Ladrón-de-Guevara, 2012). By creating narrowly defined tasks the researcher limits the crowd's intellectual and creative freedom. Symposium presenter Stuart Geiger provided the example of Galaxy Zoo (<https://www.zooniverse.org/projects/zookeeper/galaxy-zoo/>), an online citizen science project focused on astronomy that provided very specific simplified instructions for participants to code images of the galaxy. Yet, Hanny Van Arkel, a volunteer participant, found a new astronomical object that could not be identified using the given participatory structure. Only by breaking the strictures of the coding process and contacting the researchers directly could Van Arkel make the extraordinary discovery known and thus valuable to science. Geiger concluded from this that we cannot 'specify from the outset the kind of things we want our workers to do.' This anecdote contributes to the narrative that harnessing collective crowd intelligence makes big discoveries possible, while teaching that simple tasks can be creatively restricting. The idea of giving up control over knowledge production and accepting participatory input from crowd-tasker are issues that polarize researchers, evoking parallels to the controversial discussion of subjectivity in positivist and constructivist traditions. This reflects a pervasive tension in the crowdsourcing discussion between the idealized image of egalitarian and collective crowd wisdom, described by Surowiecki (2005) and Web-based peer production (Benkler, 2016) and the disaggregated, alienated mental labour typically capacitated by crowdsourcing technologies — a new digital iteration of deskilling.

Platforms as a mediator between researcher and crowd

In this section, we draw attention to the features of platforms that highly influence knowledge production that can, to a limited extent, be manipulated in designing a crowdsourcing project. The platform is the computational interface between researcher and crowd-tasker. The virtual spaces in which parties meet and thus constitute specific relations of control and power. The software design of a platform instills expectations on both sides, structures the presentation of a task, and controls the options for communication and submitting information. Platforms are created within a social context and represent the agenda and values of the designers, which in turn shapes how the software application can be used (Cooper, 2006). Thus, platform design is intrinsic to crowdsourcing projects.

Brabham cautioned that, 'it is important to be critical about what these platforms are doing: looking inside the black box of how they handle their data, what their policies are, how they eliminate users from their sets, who they don't like.' Often, the underlying structures by which crowd-tasker can access the research tasks, how tasks are distributed to the crowd, and so on, are hidden. The symposium highlighted the importance of being critical about platform design, recognizing that this mediator, for all its technical advancement, is no more neutral than traditional interfaces. Similar to research methods such as ethnographic studies, interviews, and written surveys, platforms contain biases, with the potential to produce socially responsive answering. They are no less prone to producing poor quality work than any other methods.

Yet, numerous studies that create crowdsourcing projects for the sole purpose of understanding and improving the relationship between platform design and project outcomes have shown that the design of the platform also offers researchers some control over the composition of the crowd (Behrend, *et al.*, 2011; Bücheler and Sieg, 2011; Wiggins and Crowston, 2010). A conscious and careful selection of platform parameters provides researchers with limited agency to structure the contingent crowdsourcing process and to some extent influence the crowd make-up. Key parameters are defining eligibility criteria for crowd-taskers, communicating a crowdsourcing project, selecting adequate rewards, designing the crowd-tasker interface, setting quality standards for the data, defining the number of times a task can be performed, etc. We assume that these parameters greatly affect who participates and why. For example, the variety of

motivations for participating in crowdsourcing has been highlighted in previous research, and includes both the extrinsic kind such as payment and intrinsic kind like the desire to contribute to research for its own sake (Brabham, 2012; 2008). The type of reward, moreover, complements different types of tasks. Easy and tedious tasks tend to require monetary rewards, while interesting or surprising tasks speak to intrinsically motivated crowd-taskers. In this sense, defining a reward (or other parameters of platforms) allows a researcher to adapt the crowdsourcing design according to their research goals and underlying assumptions about the crowd.

Platforms differ in the degree of pre-configuration of relevant parameters and the range of allowable customization. There are currently several types of platforms available, which we roughly divide into commercial platforms (e.g., Amazon Mechanical Turk), research-specific platforms (e.g., Zooniverse), and project specific platforms (e.g., Perseids Project), see [Table 1](#). These in turn can build on different business models. Daniel, *et al.* (2018) enumerate these as marketplace models such as AMT, contest models with rewards for the best solution such as InnoCenntive, auction models, such as Freelancer where workers bid on projects and volunteer platforms such as Galaxy Zoo. For research, the marketplace and volunteer model is the most common (Daniel, *et al.*, 2018).

Table 1: Some examples of platforms used in humanities and social sci research for crowdsourcing..		
Commercial	Research specific	Project speci
Amazon Mechanical Turk https://www.mturk.com	Zooniverse https://www.zooniverse.org/	Perseids Project https://www.perse
Figure Eight https://www.figure-eight.com	Prolific https://www.prolific.co	Text Thresher (https://bids.berke/research/text-thre)
Witkey https://www.crunchbase.com/organization/witkey#section-overview	InnoCentive https://www.innocentive.com	Open Philology https://openphilolc
Upwork https://www.upwork.com/		
Clickworker https://www.clickworker.com		
CloudFactory https://www.cloudfactory.com/		
DoMyStuff http://www.domystuff.com		
Samasource https://www.samasource.com/		

Research projects represented at the symposium included all three types. Commercial platforms offer a ready-made infrastructure requiring the least input from the researcher; however, the degree to which they can be customized to project needs is limited. Made-for-research platforms offer more flexibility but increase the complexity of set-up and management. Project specific platforms are fully customizable but require researchers to either have computational and programming skills or hire skilled professionals.

Besides software, commercial and research specific platforms provide additional services, such as recruitment of a crowd-tasker pool, communication with the crowd, and often, organizing task compensation. Each platform has its own recruitment strategy and selection and evaluation criteria for crowd-taskers; some of them even have 'community-managers' for communicating with crowd-taskers and resolving disputes and questions. A number of platforms already tap into thriving communities and thus are more effective in drawing a large crowd than others. For example, *The Adventures of Hashtag the Snail* was initially launched on a small volunteer-based platform but failed because of insufficient participation, prompting Literat to move the project to Amazon Mechanical Turk, where it flourished. The choice of platform involves balancing the burdens of customization with the benefits of the additional

service offered.

Each of these pre-set or adjustable parameters in the crowdsourcing process has strong implications for the knowledge production process and the validity of the produced knowledge. For example, Literat pointed that the decision to allow a single individual to perform one task multiple times (such as filling out a survey) may run counter to the independence of sample standards common in quantitative research. Hearst added that allowing crowd-taskers to communicate in small groups may make an answer more reliable (and thus increase research quality), whereas traditional survey research would interpret this kind of collaboration as sample contamination. The challenge is to find or create a platform that allows researchers to implement a crowdsourcing design that creates a fit between underlying assumptions about the crowd and the research goal of a particular project by configuring the parameters of the process.

Concluding discussion: Crowdsourcing as a pragmatic method

With this study we set out to explore how crowdsourcing, a process with a genesis in business, produces scientific knowledge in humanities and social sciences research. In an increasingly competitive academic climate, crowdsourcing offers researchers a cutting-edge tool for engaging with the public. This socio-technical practice's emergence from business rather than from academia may have hidden implications about the world that affect the knowledge it produces.

Our analysis reveals that translating from one arena to the other has significant consequences, which conflicts with how research has traditionally been conducted. We have showed that crowdsourcing projects can be divided up into research *about* and *with* the crowd. We have discussed three critical issues for using crowdsourcing to produce scientific knowledge: 1) the loss of control over crowd composition; 2) the task design and assumptions about a given crowd-tasker's capabilities; and, 3) the selection of a platform to host a research project. Further, that tasks can be *productive*, *reconfiguring*, or *evaluating*, where the researcher's pre-conceptions about the capabilities of the crowd suggest the tasks the crowd is capable of. Moreover, we showed how intrinsic and extrinsic rewards attract various crowds and are suitable for different tasks (intrinsic motivations suit complex an open-ended tasks and extrinsic motivations suit monotonous and closed tasks). Lastly, we discussed three types of crowdsourcing platforms: *commercial*, *research specific*, and *project specific*.

We argue that crowdsourcing requires reflection on the part of the researcher in order to ensure a strategically sound methodological design capable of producing valid data and results. The uncertain nature of the crowd and the flexibility of the method obscure the underlying epistemological assumptions of crowdsourcing (see also Marres and Weltverde, 2013). There are in-built contradictions that must be resolved: on the one hand, crowdsourcing implies that the scientific observer remains separate and distant from the subjects of observation following positivist research traditions (e.g., Popper, 1959; Kaplan, 2004); on the other hand, the process disturbs traditional knowledge hierarchies, calling attention to the context of knowledge production as in constructivist research traditions (e.g., Berger and Luckmann, 1967; Denzin and Lincoln, 2011). With other kinds of research methods — consider surveys versus qualitative interviews — researchers have well-established ideas about the capabilities and composition of research participants. Qualitative methods draw on constructivist ideas about the uniqueness and situatedness of each individual whose experiences and views of those experiences are in focus. Quantitative methods, on the other hand, build on positivist thought, where the focus is on objective knowledge and random sampling, and how society, like nature, builds on laws. In practice, crowdsourcing has been employed in studies that both seek 'absolute truth' and those that try to understand multiple constructed realities. In our data, we saw how researchers struggled to frame their research using traditional terms.

Crowdsourcing platforms mediate the relationship between researcher and crowd-tasker, connecting while also creating distance. The vague image of the crowd and attributed capabilities contribute to making crowdsourcing a method that seems suitable for both positivist and constructivist epistemological paradigms. The image of the crowd constructed by the mind of the researcher for a crowdsourcing project affect the research design, and the plan of action, all of which link the philosophical

assumptions to the method. Underlying epistemological assumptions about knowledge production thus have the potential to shape the crowdsourcing project in various ways. Crowdsourcing is not tied to one set of assumptions about the crowd-taskers. Instead, it is the researcher's implicit assumptions about the crowd that drive the methodological design. This image steers the definition of the task, the selection of a platform, the incentives offered, and so on. The underlying and implicit images a researcher has about the crowd is therefore impactful and shapes the quality and validity of knowledge produced. In sum, the great boon of crowdsourcing may be unproblematic for business but raises methodological and ethical questions for academia.

A hallmark of feminist studies is questioning the power hierarchies in all types of research, including the ways that research participants are "othered" through the building of boundaries between researcher and researched (Scantlebury, 2005). Crowdsourcing suffers more than many other methods in this regard, because of the boundaries inherent in the practical limitations related to the high number of participants and the platform design. This suggests that when employing crowdsourcing, researchers need to pay extra attention to these boundaries, which might conceal the fact that researchers and research subjects do not share a common understanding about the research and the knowledge that is created in the interaction between research participants and research project. Not all participants of a study agree to work under the conditions stipulated by the researcher. Breaches and resistances are common, which is revealed, for example, in margin notes and other scribbled information on surveys, contesting the meanings of the answers and questions (Feilzer, 2010). Or, consider the case of Van Arkel's contribution to Galaxy Zoo, discussed earlier. Researchers can choose to ignore such extra content produced by participants or they can take these into account, working out how to interpret findings. Respecting that research cannot tame or control the crowd fully also allows us to mind Wexler's warning that control sacrifices much potential that the method has to offer.

The challenge of a sound crowdsourcing project is to create a good fit between the assumed characteristics and capabilities of crowd-taskers and the research problem. The application of crowdsourcing as either a quantitative method that generates and analyzes mathematical data sets or as qualitative method that interprets meaning seemingly creates an epistemological dualism. But in reality, this bifurcation does not map as clearly on the process as assumed in theory. Rather, it calls our attention to the needs to align the research questions and underlying assumptions about who the crowd is with the selected task and other parameters of the crowdsourcing process. The importance of an adequate alignment requires great critical introspection, or what has been called reflexivity in qualitative research. Reflexivity can be defined as "thoughtful, self-aware analysis of the intersubjective dynamics of between researcher and the researched" [6]. Or, in other words, how researchers themselves affect the research process by their previous experiences and beliefs. Hence, we find that crowdsourcing does not *per se* affect meanings of scientific knowledge production, but that epistemological assumptions have the potential to shape the methodological design in every detail.

Across disciplines and paradigmatic traditions, methodologists suggest that a method has to fit the empirical world under study (Alasuutari, *et al.*, 2008). With crowdsourcing being applied across the disciplinary spectrum, thinking through this fit is particularly critical. In the pragmatist tradition the search for valid or absolute truth is less important than the adequacy of the method to answer the research question. We argue that crowdsourcing shares with pragmatism this emphasis on alignment and adequacy. This further implies that methodological design has to align with the empirical world under study. A pragmatist stance, allows researchers to group together qualitative and quantitative practices in complex mixed method designs as long as they are adequate for the defined research question (Williams and Vogt, 2011). Pragmatism has been on the rise as a frame of evaluation for what constitutes "good" research and how to think about research participants. In pragmatism, focus is on developing the most suitable procedure to answer a research question by continuously questioning, criticizing, and improving what one is doing and why, in order to reach the most appropriate (not the truest) knowledge on which to act.

Crowdsourcing is thus not only a digital method but could be considered a pragmatist method. Because of this, project investigators can draw on best practices identified in pragmatist literature when designing and executing a crowdsourcing project. We suggest that future research should more concretely work to ground crowdsourcing in pragmatism, thereby strengthening the methodological validity of this method. Finally, future

research will also need to take the social context of crowdsourcing projects more seriously, examining the conditions of knowledge production and understanding why researchers are drawn to crowdsourcing, and in that way unpack underlying assumptions about the crowd. As Lupton (2015) reminds us, the production and use of digital tools are embedded in wider political, social, and cultural processes. The promises of crowdsourcing relate to broader philosophical questions about the social world relevant to framing crowdsourcing projects methodologically.

Finally, as researchers and academic knowledge producers, we should not forget the parameters of knowledge production. We need to think about and reflect on the methodological underpinnings of new digital methods. To begin, we should reflect on who and what the “crowd” is and what this means for our particular study. To do so, we can draw on a pragmatist methodology that requires us to be candid about what we do and why, in relation to our end goal. We should remember that crowdsourcing stems from business and the structure of many commonly used platforms will shape our data. When using crowdsourcing, it is the researcher’s responsibility to reflect upon the conception of the crowd in order to achieve alignment between methodological assumptions, the research question, and the design of the crowdsourcing process.

About the authors

Lina Eklund is Assistant Professor of Human Computer Interaction at Uppsala University. Her work focuses on social interaction in and around digital technology. Current projects deal with uses on uses and practices of digital technologies in managing families, the impact of anonymity on digital sociality as well as the role of digital games in museums. Direct comments to: lina [dot] eklund [at] im [dot] uu [dot] se

Isabell Stamm is a sociologist and head of a Freigeist-Research Group on “Entrepreneurial Group Dynamics” at the Technical University in Berlin. Her research interests include interpersonal relations in groups and organizations, entrepreneurship, business-family intertwining, life courses, and cooperative research methods. E-mail: isabell [dot] stamm [at] tu-berlin [dot] de

Wanda Katja Liebermann is Assistant Professor of Architecture at Florida Atlantic University. She was trained as an architect at the University of California at Berkeley (M.Arch.) and practiced architecture in the San Francisco Bay Area for a dozen years. Her research focuses on theories and practices of architecture and urbanism in the context of the politics of disability rights and identity in the U.S. and EU. E-mail: wliebermann [at] fau [dot] edu

Acknowledgments

We are grateful to the University of California, Berkeley Humanities & Social Sciences Association for funding the symposium together with the Visiting Scholar and Postdoc Affairs Association (VSPA), Division of Arts and Humanities, Berkeley Institute for Data Science (BIDS), Center for Science, Technology, Medicine and Society as well as the participants at the symposium for taking part in this research.

Notes

1. Creswell and Plano Clark, 2007, p. 4.
2. Dewey, 1925, p. 47.
3. Small, 2011, p. 62.
4. Creswell and Plano Clark, 2007, p. 27.
5. See Snee, *et al.*, 2016, p. 229; Fielding, *et al.*, 2008.
6. Finlay and Gough, 2008, p. ix.

References

P. Alasuutari, L. Bickman, and J. Brannen (editors), 2008. *Sage handbook*

of social research methods. London: Sage.

P. Baert, 2005. *Philosophy of the social sciences: Towards pragmatism*. Cambridge: Polity Press.

T.S. Behrend, D.J. Sharek, A.W. Meade, and E.N. Wiebe 2011. "The viability of crowdsourcing for survey research," *Behavior Research Methods*, volume 43, pp. 800–813.
doi: <https://doi.org/10.3758/s13428-011-0081-0>, accessed 18 September 2019.

Y. Benkler, 2016. "Peer production and cooperation," In: J.M. Bauer and M. Latzer (editors). *Handbook on the economics of the Internet*. Cheltenham: Edward Elgar, pp. 91–119.

P.L. Berger and T. Luckmann, 1967. *The social construction of reality: A treatise in the sociology of knowledge*. Third edition. Garden City, N.Y.: Anchor Books.

H. Blumer, 1969. *Symbolic interactionism: Perspective and method*. Berkeley: University of California Press.

D.C. Brabham, 2013. *Crowdsourcing*. Cambridge, Mass.: MIT Press.

D.C. Brabham, 2012. "The myth of amateur crowds: A critical discourse analysis of crowdsourcing coverage," *Information, Communication & Society*, volume 15, number 3, pp. 394–410.
doi: <https://doi.org/10.1080/1369118X.2011.641991>, accessed 18 September 2019.

D.C. Brabham, 2008. "Crowdsourcing as a model for problem solving: An introduction and cases," *Convergence*, volume 14, number 1, pp. 75–90.
doi: <https://doi.org/10.1177/1354856507084420>, accessed 18 September 2019.

T. Bücheler and J.H. Sieg, 2011. "Understanding science 2.0: Crowdsourcing and open innovation in the scientific method," *Procedia Computer Science*, volume 7, pp. 327–329.
doi: <https://doi.org/10.1016/j.procs.2011.09.014>, accessed 18 September 2019.

G. Burrell and G. Morgan, 1979. *Sociological paradigms and organisational analysis: Elements of the sociology of corporate life*. London: Heinemann.

L. Carletti, D. McAuley, D. Price, G. Giannachi, and S. Benford, 2013. "Digital humanities and crowdsourcing: An exploration," *MW2013: Museums and the Web 2013*, at <https://mw2013.museumsandtheweb.com/paper/digital-humanities-and-crowdsourcing-an-exploration-4/>, accessed 18 September 2019.

J. Cooper, 2006. "The digital divide: The special case of gender," *Journal of Computer Assisted Learning*, volume 22, number 5, pp. 320–334.
doi: <https://doi.org/10.1111/j.1365-2729.2006.00185.x>, accessed 18 September 2019.

J.W. Creswell and V.L. Plano Clark, 2007. *Designing and conducting mixed methods research*. Thousand Oaks, Calif.: Sage.

F. Daniel, P. Kucherbaev, C. Cappiello, B. Benatalla, and M. Allahbakhsh, 2018. "Quality control in crowdsourcing: A survey of quality attributes, assessment techniques and assurance actions," *ACM Computing Surveys*, volume 51, number 1, article number 7.
doi: <https://doi.org/10.1145/3148148>, accessed 18 September 2019.

N.K. Denzin and Y.S. Lincoln (editors), 2011. *Sage handbook of qualitative research*. Fourth edition. London: Sage.

J. Dewey, 2007. *Logic: The theory of inquiry*. Plano, Texas: Saerchinger Press.

J. Dewey, 1925. *Experience and nature*. Chicago: Open Court.

L. Eklund, I. Stamm, and W. Liebermann, 2017. "Crowdsourcing as research method: Hype, hope, and hazard," *SIRG Research Reports (SIRR)*, number 1, at http://www.sirg.se/wp-content/uploads/2013/12/Crowdsourcing_SIRR_17.pdf, accessed 18 September 2019.

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

doi: <https://doi.org/10.1177/0165551512437638>, accessed 18 September 2019.

M.Y. Feilzer, 2010. "Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm," *Journal of Mixed Methods Research*, volume 4, number 1, pp. 6–16. doi: <https://doi.org/10.1177/1558689809349691>, accessed 18 September 2019.

N. Fielding, R.M. Lee, and G. Blank (editors), 2008. *Sage handbook of online research methods*. Los Angeles, Calif.: Sage.

L. Finlay and B. Gough, 2008. *Reflexivity: A practical guide for researchers in health and social sciences*. Chichester: Wiley.

D. Geiger, S. Seedorf, T. Schulze, R.C. Nickerson, and M. Schader, 2011. "Managing the crowd: Towards a taxonomy of crowdsourcing processes," *AMCIS 2011 Proceedings*, at http://aisel.aisnet.org/amcis2011_submissions, accessed 18 September 2019.

R. Glott, P. Schmidt, and R. Ghosh, 2010. "Wikipedia survey — Overview of results," at http://www.ris.org/uploadi/editor/1305050082Wikipedia_Overview_15March2010-FINAL.pdf, accessed 18 September 2019.

D.L. Hansen, P.J. Schone, D. Corey, M. Reid, and J. Gehring, 2013. "Quality control mechanisms for crowdsourcing: Peer review, arbitration, & expertise at familysearch indexing," *CSCW '13: Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, pp. 649–660. doi: <https://doi.org/10.1145/2441776.2441848>, accessed 18 September 2019.

J. Howe, 2006. "The rise of crowdsourcing," *Wired* (1 June), at <https://www.wired.com/2006/06/crowds/>, accessed 18 September 2019.

L. Irani, 2015. "The cultural work of microwork," *New Media & Society*, volume 17, number 5, pp. 720–739. doi: <https://doi.org/10.1177/1461444813511926>, accessed 18 September 2019.

L. Irani and M.S. Silberman, 2013. "Turkopticon: Interrupting worker invisibility in Amazon Mechanical Turk," *CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 611–620. doi: <https://doi.org/10.1145/2470654.2470742>, accessed 18 September 2019.

D. Kaplan (editor), 2004. *Sage handbook of quantitative methodology for the social sciences*. Thousand Oaks, Calif.: Sage.

N. Kaufmann, T. Schulze, and D. Veit, 2011. "More than fun and money. Worker motivation in crowdsourcing — A study on Mechanical Turk," *AMCIS 2011 Proceedings*, at https://aisel.aisnet.org/amcis2011_submissions/340/, accessed 18 September 2019.

A. Kittur, B. Smus, S. Khamkar, and R.E. Kraut, 2011. "CrowdForge: Crowdsourcing complex work," *UIST '11: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, pp. 43–52. doi: <https://doi.org/10.1145/2047196.2047202>, accessed 18 September 2019.

J. Kitinger, 1995. "Qualitative research: Introducing focus groups," *British Medical Journal*, volume 311, number 7000 (29 July), pp. 299–302. doi: <https://doi.org/10.1136/bmj.311.7000.299>, accessed 18 September 2019.

I. Literat, 2012. "The work of art in the age of mediated participation: Crowdsourced art and collective creativity," *International Journal of Communication*, volume 6, pp. 2,962–2,984, and at <https://ijoc.org/index.php/ijoc/article/view/1531>, accessed 18 September 2019.

D. Lupton, 2015. *Digital sociology*. London: Routledge.

T.W. Malone, R. Laubacher, and C. Dellarocas, 2009. "Harnessing crowds: Mapping the genome of collective intelligence," *MIT Sloan Research Paper*, number 4732–09, at <https://dspace.mit.edu/bitstream/handle/1721.1/66259/SSRN-id1381502.pdf>, accessed 18 September 2019.

N. Marres and E. Weltverde, 2013. "Scrapping the social? Issues in live social research," *Journal of Cultural Economy*, volume 6, number 3, pp. 313–335.

doi: <https://doi.org/10.1080/17530350.2013.772070>, accessed 18 September 2019.

A. Menking and I. Erickson, 2015. "The heart work of Wikipedia: Gendered, emotional labor in the world's largest online encyclopedia," *CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 207–210.

doi: <https://doi.org/10.1145/2702123.2702514>, accessed 18 September 2019.

J. Pedersen, D. Kocsis, A. Tripathi, A. Tarrell, A. Weerakoon, N. Tahmasbi, J. Xiong, W. Deng, O. Oh, and G.-J. de Vreede, 2013. "Conceptual foundations of crowdsourcing: A review of IS research," *2013 46th Hawaii International Conference on System Sciences*, pp. 579–588.

doi: <https://doi.org/10.1109/HICSS.2013.143>, accessed 18 September 2019.

K.R. Popper, 1959. *The logic of scientific discovery*. New York: Basic Books.

N.R. Prestopnik and K. Crowston, 2013. "Gaming for (citizen) science: Exploring motivation and data quality in the context of crowdsourced science through the design and evaluation of a social-computational system," *ESCIENCEW '11: Proceedings of the 2011 IEEE Seventh International Conference on e-Science Workshops*, pp. 28–33.

doi: <https://doi.org/10.1109/eScienceW.2011.14>, accessed 18 September 2019.

C. Robson, M. Hearst, C. Kau, and J. Pierce, 2013. "Comparing the use of social networking and traditional media channels for promoting citizen science," *CSCW '13: Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, pp. 1,463–1,468.

doi: <https://doi.org/10.1145/2441776.2441941>, accessed 18 September 2019.

R. Rogers, 2015. "Digital methods for Web research," In: R.A. Scott and S.M. Kosslyn (editors). *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource*. Hoboken, N.J.: Wiley, pp. 1–22.

doi: <https://doi.org/10.1002/9781118900772.etrds0076>, accessed 18 September 2019.

R. Rogers, 2013. *Digital methods*. Cambridge, Mass.: MIT Press.

J. Ross, L. Irani, M.S. Silberman, A. Zaldivar, and B. Tomlinson, 2010. "Who are the crowdworkers? Shifting demographics in Mechanical Turk," *CHI EA '10: CHI '10 Extended Abstracts on Human Factors in Computing Systems*, pp. 2,863–2,872.

doi: <https://doi.org/10.1145/1753846.1753873>, accessed 18 September 2019.

K. Scantlebury, 2005. "Maintaining ethical and professional relationships in large qualitative studies: A Quixotic ideal?" *>Forum: Qualitative Social Research*, volume 6, number 3, at <http://www.qualitative-research.net/index.php/fqs/article/view/35/73>, accessed 18 September 2019.

H. Shepherd, 2012. "Crowdsourcing," *Contexts*, volume 11, number 2, pp. 10–11.

doi: <https://doi.org/10.1177/1536504212446453>, accessed 18 September 2019.

H. Snee, C. Hine, Y. Morey, S. Roberts, and H. Watson (editors), 2016. *Digital methods for social science: An interdisciplinary guide to research innovation*. New York: Palgrave Macmillan.

P. Steedman, 1991. "On the relations between seeing, interpreting and knowing," In: F. Steier (editor). *Research and reflexivity*. London: Sage, pp. 53–62.

J. Surowiecki, 2005. *The wisdom of crowds*. New York: Anchor Books.

A. Tarrell, N. Tahmasbi, D. Kocsis, J. Pedersen, A. Tripathi, J. Xiong, O. Oh, and G. de Vreede, 2013. "Crowdsourcing: A snapshot of published research," *AMCIS 2013: 19th Americas Conference on Information Systems*, pp. 962–975.

M.N. Wexler, 2011. "Reconfiguring the sociology of the crowd: Exploring crowdsourcing," *International Journal of Sociology and Social Policy*, volume 31, numbers 1–2, pp. 6–20.

doi: <https://doi.org/10.1108/01443331111104779>, accessed 18

September 2019.

A. Wiggins and K. Crowston, 2010. "From conservation to crowdsourcing: A typology of citizen science," *HICSS '11: Proceedings of the 2011 44th Hawaii International Conference on System Sciences*, pp. 1–10.
doi: <https://doi.org/10.1109/HICSS.2011.207>, accessed 18 September 2019.

W. Willett, J. Heer, and M. Agrawala, 2012. "Strategies for crowdsourcing social data analysis," *CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 227–236.
doi: <https://doi.org/10.1145/2207676.2207709>, accessed 18 September 2019.

M. Williams and W.P. Vogt (editors), 2011. *Sage handbook of innovation in social research methods*. Los Angeles, Calif.: Sage.

J. Zittrain, 2008. *The future of the Internet and how to stop it*. New Haven, Conn.: Yale University Press.

Editorial history

Received 4 June 2019; accepted 28 August 2019.

Copyright © 2019, Lina Eklund, Isabell Stamm, and Wanda Katja Liebermann. All Rights Reserved.

The crowd in crowdsourcing: Crowdsourcing as a pragmatic research method
by Lina Eklund, Isabell Stamm, and Wanda Katja Liebermann.
First Monday, Volume 24, Number 10 - 7 October 2019
<https://journals.uic.edu/ojs/index.php/fm/article/view/9206/8124>
doi: <http://dx.doi.org/10.5210/fm.v24i10.9206>

A Great Cities Initiative of the University of Illinois at Chicago [University Library](#).

© *First Monday*, 1995-2020. ISSN 1396-0466.