Stockholms universitet

Stockholm Internet Research Group
Dept. of Sociology /www.sirg.se

CROWDSOURCING AS RESEARCH METHOD: HYPE, HOPE, AND HAZARD

Lina Eklund[1], Isabell Stamm[2] and Wanda Liebermann[3]
[1]Uppsala University, [1]SIRG, [2]Technical University in Berlin, [3]Florida Atlantic University

SIRG
Research Reports
SIRR, 2017:1

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

Keywords: *Crowdsourcing, Research Method, Digital Method, Platforms, Pragmatism*

Abstract

This study aims to untangle crowdsourcing as a research method by drawing attention to the ways in which crowdsourcing affects meanings of scientific knowledge production. Our analysis is inspired by the proceedings of a symposium on crowdsourcing in the humanities and social science held at UC Berkeley in 2015, we then draw on a literature review in conjunction with content analysis of ten websites of crowdsourcing service providers. Using these sources, we perform a theorizing analysis on the practicalities of this digital method. We show that the translation of crowdsourcing from a business process to research method is not trivial. This sociotechnical practice rests upon an unpredictable composition of the crowd, malleability of tasks, and in-built ambiguity of platforms; which leaves room to contradictory images of the crowd. Which we argue produces problems for scientific knowledge production in the traditional research epistemologies of positivism and constructivism. The ambiguity in what a crowd is, while making crowdsourcing adaptable to quantitative, qualitative, and mixed-method designs, can be problematic when producing scientific knowledge. We conclude that its adaptability distinguishes crowdsourcing as a pragmatist method that requires great reflexivity on behalf of researchers.

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

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, November 6, 2016

A large group of researchers gathered for a conference on the UC Berkeley campus are listening to a presentation and waiting for their turn to view the children's book 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 artifact—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 as complicated as a book then the potential for research application seems vast. Indeed, as 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). In essence, crowdsourcing harnesses the time, energy, and talents of individuals, whom we call crowd-taskers[1], reached through the Internet to perform a given undertaking (Shepherd, 2012). Its hybrid etymology—"crowd" plus "outsourcing"—signals

[1] 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.

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

ambiguous potentials: from the exacerbation of dispersed, part-time, tedious, piecework-based labor conditions enabled through web platforms (Irani, 2015) to, more hopefully, 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.

The novelty of this technique of research has yet to produce routines and best practices, and more importantly consensus is lacking—much less among disciplines—about what this "method" is, how to use it, and even why to use it. The preponderance 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; Prestopnik and Crowston, 2013), and recruitment and retention of participants (Prestopnik and Crowston, 2013; Robson et al., 2013). Yet, the particular cultural status of scholarly knowledge, with defined criteria for establishing valid knowledge, distinguishes an analysis of crowdsourcing as a scientific research method from its uses in commercial, civic, and other sectors. To date, there is scant critical scholarship on crowdsourcing as a research method that considers its epistemological implications (Pedersen et al., 2013).

Crowdsourcing is not native to research and its origin in business comes with certain hidden assumptions about the world which concretely affect the knowledge produced. We thus ask, can crowdsourcing still qualify as scientific method and how can we understand its methodological underpinnings? In this study, we aim to demystify crowdsourcing as used in humanities and social science research and connect that discussion to a broader debate about scientific knowledge production. This link is not trivial 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 foreground important methodological implications of key ideas circulating about crowdsourcing.

Methodology

Our methodology consists of a theorizing process (Swedberg, 2012) where we combine various types of empirical data which we then theorize in order to shed light on our topic. First, we participated in a symposium, detailed below, in order to investigate uses of crowdsourcing in academia; we then studied crowdsourcing web-sites in order to investigating the promises of the method, and finally we studied the academic literature on crowdsourcing.

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

The starting point of our study was a symposium on scientific crowdsourcing held at UC Berkeley in 2016, "Crowdsourcing and the Academy: Exploring promises and problems of collective intelligence methods in humanities and social science research". 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 that followed a 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, and "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 4200 examples of written and visual material—highlights the breadth of possibilities for Internet-facilitated creativity. Next was the Perseids Project, 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 classrooms and the general public, grants contributor authorship, and uses a complex review and feedback system. Third, The Deciding Force Project, led by Nick Adams of the Berkeley Institute for Data Science (BIDS), produced its own software called Textthresher, 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 analyzes an enormous dataset of newspaper articles chronicling police and Occupy Movement interaction. Finally, UC Berkeley School of Information's Marti Hearst presented the project, Improving Crowdwork with Small Group Discussions, offering insights into how to facilitate peer evaluation and co-learning among crowd workers in order to enable more challenging tasks. The projects do not represent the full breadth of crowdsourcing in research but, rather, each highlights 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 UC 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

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

sharing practices based on digital technologies. Finally, Stuart Geiger of UC Berkeley's School for 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: keynote speech, panel presentations, and panel and audience discussion, which were recorded with the permission of the organizers and presenters.

Additionally, we screened the websites of ten crowdsourcing platforms and service providers commonly used in research—Amazon Mechanical Turk, Clickworker, Cloudworking, Crowdflower, Crowdsource, Crowdsourcing Week, Innocentive, Microworkers, Workhub, and Zooniverse—for representations of the crowdsourcing process. The aim was to investigate the promises of the method. We thus only looked at the front end of the sites, the information the pages give to attract new projects. The platforms were chosen based on two criteria: they were mentioned during the symposium, and they are websites that cater specifically to researchers. Text and images from the web-sites where coded inductively by the promises of this method.

Finally, we performed a literature search (Webster and Watson, 2002) of scholarly publications on the topic of crowdsourcing in the social sciences and humanities. The aim was to investigate the current status of crowdsourcing as a scientific method and through this map out future directions for the use and practice of this method (Webster and Watson, 2002). We started with the publications of the symposium presenters and worked outwards. We read more than a hundred articles and books among the three authors, wrote summaries and met to discuss these summaries. Key works were read among the authors and then discussed in-depth. Surprisingly, very few works dealt with the methodological implications of crowdsourcing or the evaluation of crowdsourcing used in research. The bulk of academic publishing on crowdsourcing pertains to who the crowd-taskers are and their motivations, together with discussions about ethics and implication for work this type of labor holds (e.g., see the work of Lily Irani).

Our study is a review of the current field of academic crowdsourcing and as Webster and Watson (2002) argue, theorizing is crucial for any useful review. We thus draw on Swedberg (2012) in order to theorize crowdsourcing as a scientific method. Theorizing (Swedberg 2012) should not be confused with presenting a readymade theory. Theorizing is what happens in the context of discovery. We started with empirical facts from three venues of data gathering and allowed these to drive our analysis (Swedberg, 2012). The end result was an analytical process inspired by grounded theory where we used topic categorization to find common themes and trends in our material. The interview transcripts, summaries and screenshots from the

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

web-providers, and summaries from the literature review made up a data-base that we scanned manually in our analysis. Our topics where refined and reworked until all three authors where in agreement; we thus did many rounds of writing, thinking, and talking through our data. We identified five tropes—or hopes—about crowdsourcing as a research method, presented below, and three key areas for theorizing about this emerging method: 1) crowd-taskers as both subject and objects, 2) research tasks 3) the influence of platforms. An additional key area was identified, namely the ethical implications of this method. We had to exclude ethics from the article at hand as the topic was broad enough to grant its own separate investigation.

Crowdsourcing as a digital method

A research method is a set of techniques or heuristics that define a system for studying a phenomenon. Research methods vary depending on the traditions of the subject field, creating a framework for a particular viewpoint (Alasuutari et al., 2008). 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. It can thus be understood as a research method (Snee et al., 2016). More specifically, crowdsourcing is a digital method in that it uses online and digital technologies to collect and analyze 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 online. This contrasts with methods that have become digitized, such as online surveys or online ethnography, traditional methods that have migrated online.

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 Magazine, reminds us that crowdsourcing is a tool originally developed for commercial purposes, which launched businesses like iStockphoto.com and threadless.com. Marres and Weltevrede (2013) caution us about the translation of crowdsourcing into research context, observing that its application in academia requires critical reflection on the built-in values inherited from a commercial process and thus transmitted to the data it generates.

Central to defining crowdsourcing as a method for research is thus an understanding of how it fits into research design. Research design is commonly thought of as a 'plan of action that links philosophical assumptions to specific method' (Creswell and Plano Clark, 2007: 4). According to Blumer (1969) a research design includes developing a picture of the empirical world, asking questions about

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

that world and turning these into researchable problems, and finding the best means of doing so, which involves making choices about methods, the development and use of concepts, and the interpretation of findings (Blumer 1969).

For digitized research methods like online surveys or online ethnography, researchers have successfully revised their underlying assumptions of knowledge production (see Snee et al., 2016: 229; Fielding et al., 2008). Yet, we have just begun to probe the philosophical framework underlying the foundations of methods such as crowdsourcing (Rogers, 2013). As crowdsourcing continues to penetrate academic research it becomes increasingly important to recognize that, as with any methodology, crowdsourcing carries specific assumptions of the knowledge produced.

If crowdsourcing is a digital method, we need to understand the underlying philosophical assumptions of this method. To do so we need to comprehend and chart underlying epistemologies. Two key research paradigms have dominated historically: positivism (realism) and constructivism (relativism). Paradigms here represents established standards, certain way of viewing the world (Burell 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 behavior. Constructivism tend to favor humanistic qualitative methods and argue that people experience and understand the same 'objective reality' in very different ways and have their own, often very different, reasons for acting in the world. These two paradigms have at times been at odds as they advance 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. Recently, pragmatism has gained traction as a research philosophy which could end this rift between the old qualitative versus quantitative or data-driven versus interpretative methods debate (Baert 2005). Pragmatism as a research philosophy aims "to set aside considerations about what is ultimately true in favor of what is ultimately useful." (Small 2011: 62). Pragmatism acknowledge how our method of inquiry may alter the scientific knowledge produced; in a sense, knowledge is seen as active. Method, in a pragmatic sense, always entails the question of aim, and thus partially depends on what a researcher wants to achieve (Baert 2005). Pragmatism has been particularly influential in the mixed methods tradition as it has this potential to bridge between different research traditions. Using a pragmatic approach, we can see methods as active agents in shaping knowledge, and that there are is no single underlying assumption that defines a 'correct' method. Rather, depending on what we want to achieve, different methods will be useful.

We will now move onto examining the promises of crowdsourcing as an academic method as this will shed light on the process and then we will discuss key questions

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

about the method, and lastly, end with a discussion on the philosophical underpinnings of this digital method.

Promises of crowdsourcing as an academic method

There are many practical and ideological reasons that explain why researchers are drawn to a process whose origins are in labor outsourcing. We identified five promises, or tropes, within the crowdsourcing narrative.

1) The efficiency trope: Crowdsourcing platforms have a tendency to present crowd work as a "magical" process: the researcher inputs a task, and enriched data comes back. Crowdflower's website (crowdflower.com) uses the image of a launching rocket that returns with results, invoking speed, technological innovation, and a journey into the unknown (Figure 1). Along the same lines, Mechanical Turk promises that the researcher will 'start receiving results in minutes' (mturk.com), while Workhub offers ways to 'use the Internet to get a year's work done in a day.'[2]

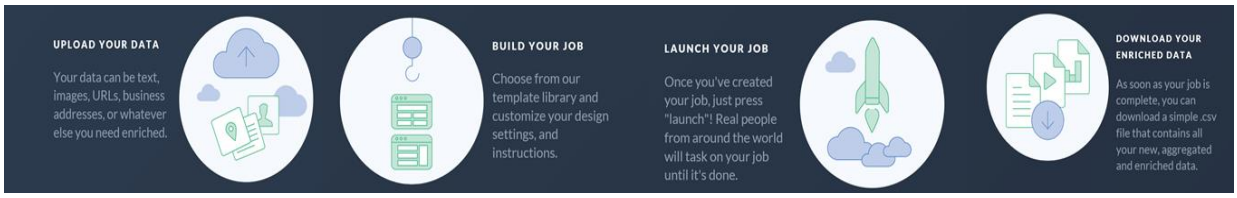


Figure 1: The process of crowdsourcing as illustrated on crowdflower.com

2) The scaling up trope: Crowdsourcing promises to scale up research both in terms of the number of participants and quantity of data. While interest is intensifying among scholars to capture the vast quantity of data on the web, they struggle with how to do it. As Nick Adams said at the symposium, 'There is all this digital text out there; it is pretty cheap to acquire; […] There are terra bytes of textual data that come online every day.' Crowdsourcing is seen as providing access to 'big data', because it employs the crowd for a wide array of tasks, like taking pictures, writing text, and recording stories, and to digest web based data such as tweets, posts, and links. The capacity to process large and unique datasets creates advantages for researchers competing in an environment that values novelty, the current benchmark for publication and research grants. Digitally harnessing collective human intelligence

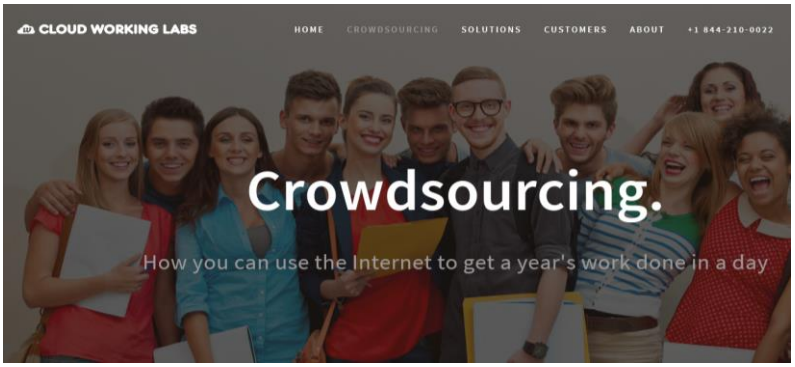
[2] Work hub has since gone out of business.

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

appears to offer a way to handle big data while overcoming the limitations of algorithms and computational methods for analysis. Karpf, who observes that data abundance is a 'mirage' due to its susceptibility to the rapid disappearance of data and obsolescence of platforms over time (2012), is a notable exception to the uncritical advocacy.

3) The unlimited workforce trope: Crowdsourcing platforms build on the entrepreneurial image of crowd work—flexible, self-directed, project-based, market oriented, and geographically unconstrained. At the symposium, Marti Hearst stated that 'a lot of people think [that crowdsourcing] may transform work for everyone in the future.' Images of young and diverse professionals promise clients an accommodating global labor pool, always at the ready (Figure 2). Similarly, Mechanical Turk offers 'access to an on-demand, scalable workforce' (mturk.com), and Crowdsource provides employers a 'skilled and scalable workforce' (crowdsource.com). As Lily Irani noted,

> One of the things that [crowdsourcing] enables is, for somebody who has a bunch of data processing to do in a pinch […] can hire 10,000 workers for a couple of days and just blow through that; so, it makes labor relations approach a scale where you have one manager interacting with, say, 10,000 workers, for a very short time' (2015).



© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

Figure 2: The crowd as depicted on the websites of crowdsourcing.org (above) and microworkers.com (below)

Using the word 'workforce' gives the impression of a big company with many employees, when in reality these sites have 'registered users,' whose skill level is often untested. Consequently, a main area of the crowdsourcing literature examines ways to increase standards for participant performance, such as quality control measures (Hansen et al., 2013) and recruitment and retention of participants (Robson et al., 2013; Prestopnik and Crowston, 2013).

(4) The creative community trope: Crowdsourcing promises voluntary and creative collective action, a common narrative in Web 2.0 discourse. For example, Cloudwork announces on its website that, 'Crowdsourcing is what happens when you pool the resources of hundreds, thousands, or millions of people towards a common goal' (cloudwork.com). What arises is a 'solver community' (innocentive.com) devoted to 'people powered research' (zooniverse.org). Brabham points out that this strips out the "hacker" ethos of the web leaving only a benign and constructive image (Brabham, 2008, p. 82). Wexler (2011), in a rare social theory critique of crowdsourcing, historicizes today's crowd rhetoric by connecting it to earlier debates that arose alongside modern democratic states and the elites' fears of mob control. He writes that, 'The contemporary literature on crowdsourcing casts the classical social scientific conception of the crowd from social problem to problem solver' (Wexler, 2011). The faceless crowd goes from uncontrollable, to controlled.

(5) The dissolving intersubjectivity trope: Besides efficiency, the rocket metaphor used by crowdflower.com also suggests a desirable opacity between researcher and contributors: they live on separate planets, too distant to have a clear picture of each other. Commercial sites perpetuate a dream of effortless efficiency through narratives that portray crowdsourcing for research as harnessing the full potential of Web 2.0 technologies, the new wave of the Internet characterized by user-generated content and social media. Crowd and cloud discourse share the digital

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

imaginary of an ethereal, frictionless realm where contact points, labor, and other actions are removed from their earthly material context (Hu, 2015). This vision transforms the exertion of face-to-face fieldwork to a virtual space where researcher and crowd engage through a mediating online-platform.

In sum, crowdsourcing is envisioned as a technological innovation capable of transforming research through amplified processing power and by creating limitless communication circuits between researcher and previously unimaginable numbers of contributors. Equally captivating is the crowdsourcing rhetoric of unleashing a vast creative potential, which will enable remarkable discoveries while at the same time educating and building a knowledge community. Underlying these tropes is an image of the crowd as a faceless, idealized entity, rather than a group of individuals. This nebulous crowd identity is enhanced by the mediating role of the platform, which makes room for all sorts of images about who crowd-taskers are and what they can do. It is clear that crowdsourcing's origin in the business sector affect the potential and capabilities represented in the description of this method. Crowdsourcing is not native to research and its origin thus comes with certain assumptions about the world, which potentially affect the scientific knowledge produced by this method. A question we turn to now.

Critical issues in the crowdsourcing process

Crowd-taskers as subject and object of crowdsourced research

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, which differs from civic, commercial, and artistic applications. Likely because of this nebulous character of the crowd, research on the socio-demographic composition of the crowd and the working conditions of the individual crowd-taskers has proliferated. In other words, research *about* the crowd. Understanding the crowd informs research *with* the crowd both as *object* of research, i.e. as a specific group that informs our understanding of the social world, and as *subject* in research, i.e. as capable individuals assisting in the process of knowledge production.

Studies *about* the crowd have been attempting to dispel 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. As shown above. However, the digital divide still exists. Many areas of the world have inadequate or no access to the

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

Internet; and even in countries with almost universal access, some individuals will be more Internet savvy than others. As we have shown, crowdsourcing platforms also play on the image of a vast, diverse, and global crowd in their marketing. In reality the crowd is far less diverse than often assumed, as studies on the identity of crowd workers and their working conditions have shown (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).

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 Salt Lake City, Utah bus stop shelter design project that Brabham presented was designed with a local Utah audience in mind, open to anyone willing to participate. While the Perseids project succeeded in attracting the aspired crowd, the bus stop competition organizers were 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, as crowdsourcing projects are communicated online, who will accept a task is unpredictable. We thus argue that determining a universal definition of the crowd is not possible as each project design instantiates a unique composition of crowd-taskers, which cannot be entirely predicted beforehand.

This circumstance has implications for doing research *with* the crowd. The ambiguity about who the crowd is allows for a great deal of uncertainty about the nature of the crowd; which creates unique difficulties for this method. In other types of research, we have a measure of control of who is in our study: we use controlled random samples or we meet our informants face-to-face. Due to the anonymous nature

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

of the Internet, these control mechanisms are put out of play. Instead, we run the risk that tropes about the crowd create proxy meanings that affect project design.

Adhering to efficiency, effortless intersubjectivity, or/and scaling up tropes implies an understanding of the crowd as research object. These types of studies depend on scientific rules of representativeness, which conceptualize the crowd as a study sample. Interestingly, research on crowdsourcing has found that crowdsourcing 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. The unpredictability and homogeneity of the crowd challenge scientific rules of representativeness, 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 community members. As we have shown, however, crowdsourcing cannot be limited to this one version. The differing underlying assumptions about who the crowd is and reduction of research participants into a crowd go along with epistemological assumptions about the relationship between knower and the known, and hierarchies and power relations in knowledge production.

Research tasks reflect 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), the illustration of a children's book (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. First, *productive tasks* involve the generation of

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

ideas, designs, data, or text, which includes 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 a higher order concept 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) makes 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 the 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/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 likely to vary. These problems lessen the researcher's control over results and undermine standards of good research. For the Deciding Force Project, Adams approached these risks with two main simplifying strategies. First, as other authors have noted is common to crowdsourcing projects (e.g. Estellés-Arolas and González-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' (Nick Adams, 2015). In this way Adams increases the comparability and standardization of results, and pulls the locus of control back to the researcher.

However, researchers who try to develop techniques for enabling crowd creativity through more complex tasks (Kittur et al., 2011) view such simplification

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

strategies as 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, 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. However, the idea of giving up control over knowledge production and accepting participatory input from crowd-tasker are issues that appear to 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, forthcoming), and the disaggregated, alienated mental labor typically capacitated by crowdsourcing technologies—a new digital iteration of deskilling.

Indeed, the great draw of crowdsourcing is the method's ability to draw on large numbers of individuals. Because of the complexities of managing huge numbers of persons, 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. This complexity-reducing mechanism is seen as the great benefit of crowdsourcing for business, yet becomes inherently problematic when applied to research, as it contradicts the basic idea that we control who participates in our studies, either as part of our sample or as part of our team.

Platforms as a mediator between researcher and crowd

In this section, we draw attention to the features of platforms that highly influence knowledge production which can to a limited extend 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

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

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, however, 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 discussion 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 like ethnographic studies, interviews, and written surveys, platforms contain biases, with the potential to produce skewed answers. They are no less prone to producing poor quality work than other methods.

Yet, as has been shown in numerous studies that create crowdsourcing projects for the sole purpose of understanding and improving the relationship between platform design and project outcomes, 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 extend influence the crowd make-up. Key parameters are defining eligibility criteria for crowd-taskers, communicating a crowd-sourcing 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 can assume that these parameters greatly affect who participates and why. For example, the variety of motivations for participating in crowdsourcing have been highlighted in previous research, and include both the extrinsic kind such as payment and intrinsic kind like the desire to contribute to research for its own sake (Brabham, 2008, 2012). The type of reward, however, 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 platform parameters) allows a researcher to adapt the crowdsourcing design according to their research goals and underlying assumptions of the crowd.

Platforms differ in the degree of pre-configuration of relevant parameters and the range of customization they allow for. 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

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

platforms (e.g. Perseids Project). The 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, like 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. Deciding which kind of platform involves balancing the burdens of customization with the benefits of the additional service offered by platforms.

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. Literate pointed out that some decisions, for example, that a single individual can perform one task multiple times (such as filling out a survey), may run counter to the independence of sample standard common in quantitative research. And Hearst suggested that allowing crow-taskers to communicate in small groups may make an answer more reliable (and thus increase research quality). The challenge is to find or create a platform that allows a research project to implement a crowdsourcing design that creates a fit between underlying epistemological assumptions about the crowd and the research goal of a particular project by configuring the parameters of the process.

Concluding discussion: Crowdsourcing, 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. Yet this socio-technical practice emerged as a business procedure rather than as a research method and thus

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

contains many hidden assumptions about the world which concretely affect the knowledge produced.

Our analysis suggests that changing arenas has significant implications, for example, applying crowdsourcing in research retains certain intrinsic values from the business realm, which often conflict with academic research standards. We have shown that a crowdsourced project can be divided up in research *about* and *with* the crowd. Further, that tasks can be productive, reconfiguring or evaluating, and that researcher's pre-conceptions about the capabilities of the crowd inherently suggest the types of tasks the crowd is capable of. Moreover, we showed how intrinsic and extrinsic rewards attract various crowds and are suitable for different tasks (intrinsic rewards suit complex an open ended tasks and extrinsic rewards suit monotonous and closed tasks). Lastly, we discussed three types of crowdsourcing platforms: commercial, research specific, and project-specific. We then discussed three critical issues which we identified in the theorizing process: the nebulous concept of the crowd, malleability of tasks, and in-built biases of platforms developed for business. Crowdsourcing thus requires reflection on the part of the researcher in order to ensure a strategically sound methodological design capable of producing valid data and interpretation. The vague image of the crowd and the malleability of the method obscure the underlying epistemological assumptions of crowdsourcing (Marres and Weltverde, 2013). On the one hand, crowdsourcing implies that the scientific observer remains separate and distant from the entities that are subject 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—just think about surveys vs. qualitative interviews—we have well-established ideas about the capabilities and composition of our research participants. In qualitative methods, we draw on constructivist ideas about the uniqueness and situatedness of each individual whose experiences and views of those experiences are in focus. On the other hand, quantitative methods build on positivist thought, where the focus is on objective knowledge and random sampling, and how society, like nature, builds on absolute laws. Hence, it remains unclear whether crowdsourcing can be used to search for 'absolute truth', or to understand multiple constructed realities, or whether it values what is ultimately useful eschewing both standpoints. The latter is close to pragmatism (Dewey, 2007).

Our analysis of the promises of crowdsourcing revealed five narrative tropes about this sociotechnical practice. Most striking, these promises all contain deliberately vague characterizations of the crowd, perpetuated by the platforms, which creates an obscuring interface between the crowd and researcher. Crowdsourcing

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

platforms mediate the relationship between researcher and crowd-tasker, connecting while also creating distance. The vague images of crowds and their assumed capabilities contribute to making crowdsourcing a method that seems to fit with either positivist or constructivist epistemological paradigms. However, which image we have of the crowd will work as a base for a crowdsourcing project and affect our research design, our plan of action which links our philosophical assumptions to our method. Underlying epistemological assumptions about knowledge production can thus have the potential to shape the crowdsourcing project in various ways. Crowdsourcing, however, 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 implied images a researcher has about the crowd is thus impactful and shapes the quality and validity of knowledge produced. Thus, the great boon of crowdsourcing may be unproblematic for business, but raises methodological and ethical questions for academia.

The challenge of a sound crowdsourcing project is thus to create a fit between the assumed characteristics and capabilities of crowd-taskers and the research problem. The application of crowdsourcing as both either a quantitative method that generates and analysis mathematical data sets or as qualitative method that interprets meaning seemingly creates an epistemological dualism, which in reality does not map as clearly on the process as theoretically assumed. Rather, it calls our attention the inscribed 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. Hence, we find that crowdsourcing does not per se affect meanings of scientific knowledge production—but the applied image of the crowd has the potential to shape the methodological design in all its 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 not trivial. In pragmatist traditions search for valid or absolute truth is less important as the adequacy of the method for the question. We argue that crowdsourcing shares with pragmatism this emphasis on alignment and adequacy of the method for the question. This further implies that our 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

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

constitutes “good” research and how to think about research participants. In pragmatism, focus is on reaching the most suitable procedure to answer a research question by constantly questioning, criticizing, and improving what one is doing and why, in order to reach the most appropriate (note: not the truest) knowledge on which to act.

Crowdsourcing is thus not only a digital method, but also a pragmatist method. Because of this, we can draw on best practices identified in pragmatist literature when we design and execute a crowdsourcing project. We suggest that future research should more concretely work to ground crowdsourcing in pragmatism; thus 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 thereby unpacking underlying assumptions about the crowd. As Lupton (2015) reminds us, the production and use of digital tools are embedded in a wider political, social, and cultural process. The promises of crowdsourcing relate to broader philosophical questions about our social world relevant to frame crowdsourcing projects methodologically.

Finally, 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 image of the crowd in order to achieve alignment between methodological assumptions, the research question, and the design of the crowdsourcing process.

References

Alasuutari P, Bickman L and Brannen J (2008) *The SAGE Handbook of Social Research Methods*. London: SAGE.

Baert, P.(2005) *Philosophy of the Social Sciences: Towards Pragmatism*. 1 edition. Polity.

Behrend TS, Sharek DJ, Meade AW and Wiebe EN (2011) The Viability of Crowdsourcing for Survey Research. *Behavior Research Methods* 43(3): 800–813.

Benkler Y (forthcoming) Peer Production and Cooperation. In: Bauer JM and Latzer M (eds) *Handbook on the Economics of the Internet*. Cheltenham: Edward Elgar.

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

Berger PL and Luckmann T (1967) *The social construction of reality*. New York: Anchor Books.

Blumer H (1969) *Symbolic Interactionism: Perspective and Method*. Berkeley: University of California Press.

Brabham DC (2008) Crowdsourcing as a Model for Problem Solving An Introduction and Cases. *Convergence: The International Journal of Research into New Media Technologies* 14(1): 75–90.

Brabham DC (2012) The Myth of Amateur Crowds. *Information, Communication & Society* 15(3): 394–410.

Brabham DC (2013) *Crowdsourcing*. Cambridge, MA: MIT Press.

Bryman A (2008) The End of Paradigm Wars? In: Alasuutari P, Bickman L and Brannen J (eds) *The SAGE Handbook of Social Research Methods*. London: SAGE, pp. 13–25.

Bücheler T and Sieg JH (2011) Understanding Science 2.0: Crowdsourcing and Open Innovation in the Scientific Method. *Proceedings of the 2nd European Future Technologies Conference and Exhibition (FET)* 7: 327–329.

Burrell, G., & Morgan, G. (1979). Two dimensions: Four paradigms. Sociological paradigms and organizational analysis, 21-37.

Carletti LG, Price GD, McAuley D and Benford S (2013) *Digital Humanities and Crowdsourcing: An Exploration*. Available at: https://ore.exeter.ac.uk/repository/handle/10871/17763.

Cooper J (2006) The Digital Divide: The Special Case of Gender. *Journal of Computer Assisted Learning* 22(5): 320–334.

Creswell JW and Plano Clark VL (2007) *Designing and Conducting Mixed Methods Research*. Thousand Oaks, Calif.: Sage Publications.

Denzin NK and Lincoln YS (2011) *The Sage handbook of qualitative research*. Thousand Oaks: SAGE.

Dewey J (2007) *Logic: The theory of inquiry*: Saerchinger Press.

Estellés-Arolas E and González-Ladrón-de-Guevara F (2012) Towards an Integrated Crowdsourcing Definition. *Journal of Information Science* 38(2): 189–200.

Fielding N, Lee RM and Blank G (2008) *The SAGE Handbook of Online Research Methods*. Los Angeles, London: SAGE.

Geiger D, Seedorf S, Schulze T, Nickerson RC and Schader M (2011) Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes. *AMCIS 2011 Proceedings - All Submissions*.

Hansen DL, Schone PJ, Corey D, Reid M and Gehring J (2013) Quality Control Mechanisms for Crowdsourcing: Peer Review, Arbitration, & Expertise at Family Search Indexing. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*: 646–660.

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

Howe J (2006) The rise of crowdsourcing. *Wired Magazine*, 2006.
Irani L (2015) The Cultural Work of Microwork. *New Media & Society* 17(5): 720–739.
Irani L and Silberman MS (2013) Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. *Proceedings of the CHI 2013* Paris, April 27-May 02: 611–620.
Kaplan D (2004) *The SAGE handbook of quantitative methodology for the social sciences*. Thousand Oaks, Calif., London: SAGE.
Karpf D (2012) Social Science Research Methods in Internet Time. *Information, Communication & Society* 15(5): 639–661.
Kittur A, Smus B, Khamkar S and Kraut RE (2011) *CrowdForge: Crowdsourcing Complex Work*. Santa Barbara, CA.
Kitzinger J (1995) Qualitative Research. Introducing Focus Groups. *British medical journal* 311(7000): 299.
Literat I (2012) The Work of Art in the Age of Mediated Participation: Crowdsourced Art and Collective Creativity. *International Journal of Communication* 6: 23.
Lupton D (2015) *Digital Sociology*. London, New York: Routledge.
Malone T, Laubacher R and Dellarocas C (2009) *Harnessing Crowds: Mapping the Genome of Collective Intelligence*. Cambridge, MA.
Marres N and Weltverde E (2013) Scrapping the Social? Issues in Live Social Research. *Journal of Cultural Economy* 6(3): 313–335.
Pedersen J, Kocsis D, Tripathi A, Tarrell A, Weerakoon A, Tahmasbi N, et al. (2013) Conceptual Foundations of Crowdsourcing: A Review of IS Research. *46th Hawaii International Conference on System Sciences (HICSS)*: 579–588.
Popper KR (1959) *The Logic of Scientific Discovery*. New York: Routledge.
Prestopnik NR and Crowston K (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. *IEEE eScience Conference* Stockholm.
Robson C, Hearst M, Kau C and Pierce J (2013) Comparing the Use of Social Networking and Traditional Media Channels for Promoting Citizen Science. *Proceedings of the 2013 Conference 2013*: 1463–1468.
Rogers R (2013) *Digital Methods*. Cambridge, Massachusetts: The MIT Press.
Rogers R (2015) Digital Methods for Web Research. In: Scott RA and Kosslyn SM (eds) *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource*. Hoboken, N.J.: John Wiley & Sons, pp. 1–22.
Ross J, Irani L, Silberman MS, Zaldivar A and Tomlinson B (2010) Who Are the Crowdworkers?: Shifting Demographics in Mechanical Turk. *CHI '10 Extended Abstracts on Human Factors in Computing Systems* New York: ACM: 2863–2872.

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.

Shepherd H (2012) Crowdsourcing. *Contexts* 11(2): 10–11.

Snee H, Hine C, Morey Y, Roberts S and Watson H (2016) *Digital Methods for Social Science: An interdisciplinary guide to research innovation*. Houndmills, Basingstoke, Hampshire, New York, NY: Palgrave Macmillan.

Surowiecki J (2005) *The Wisdom of Crowds*. New York: Anchor Books.

Swedberg, R. (2012). Theorizing in sociology and social science: Turning to the context of discovery. *Theory and society*, 41(1), 1-40.

Tarrell A, Tahmasbi N, Kocsis D, Pedersen J, Tripathi A, Xiong J, et al. (2013) *Crowdsourcing: A Snapshot of Published Research*. Chicago, Illinois.

Wexler MN (2011) Reconfiguring the Sociology of the Crowd: Exploring Crowdsourcing. *International Journal of Sociology and Social Policy* 31(1/2): 6–20.

Wiggins A and Crowston K (2010) From Conservation to Crowdsourcing: A Typology of Citizen Science. *44th Hawaii International Conference on System Sciences (HICSS)*: 1–10.

Willett W, Heer J and Agrawala M (2012) Strategies for Crowdsourcing Social Data Analysis. *Proceedings of the SIGCHI. New York: ACM*: 227–236.

Williams M and Vogt WP (2011) *The SAGE Handbook of Innovation in Social Research Methods*. Los Angeles: SAGE.

Zittrain J (2008) *The Future of the Internet and How to Stop it*. New Haven: Yale University Press & Penguin UK.

© Copyright is held by the author(s). SIRRs receive only limited review. Views and opinions expressed in SIRRs are attributable to the authors and do not necessarily reflect those held at the Department of Sociology.