

# DON FALLIS

---

## INTRODUCTION: THE EPISTEMOLOGY OF MASS COLLABORATION

Human beings regularly work together to get things done. In particular, people frequently collaborate on the production and dissemination of knowledge. For example, scientists often work together in teams to make new discoveries. How such collaborations produce knowledge, and how well they produce knowledge, are important questions for epistemology. In fact, several epistemologists (e.g., Hardwig 1991, Thagard 1997, Wray 2002) have addressed such questions regarding collaborative scientific research.<sup>1</sup>

While most collaborations involve only a few people, new information technologies now allow huge numbers of people (separated by very large distances) to work together on a single project. For example, thousands of programmers have collaborated on open source software projects, such as the *GNU/Linux* operating system (Duguid 2006). Other notable mass collaborations include:

- *Wikipedia* ([www.wikipedia.org](http://www.wikipedia.org)), the “free online encyclopedia that anyone can edit.”
- *Yahoo! Answers* ([answers.yahoo.com](http://answers.yahoo.com)), which also allows you to get an answer to just about any question by leveraging the collective wisdom of many other Internet users. (*ChaCha* provides a similar service over your cellphone.)
- The *Great Internet Mersenne Prime Search* ([www.mersenne.org](http://www.mersenne.org)), which allows you to use your personal computer to aid in a large collective search for very large prime numbers.
- *Digg.com*, which allows you to find important and interesting news stories by collecting together those stories that have been highly ranked by many other Internet users.
- *SETI@home* ([setiathome.berkeley.edu](http://setiathome.berkeley.edu)), which allows you to participate in a large collective attempt to find evidence of extraterrestrial life by analyzing radio signals with your personal computer.

When such collaborations take place over the Internet, they are typically referred to as *Web 2.0* projects. While not all Web 2.0 projects have the goal of producing and disseminating knowledge, many of them do. Moreover, large numbers of people are now participating in such projects, and even more people are using these projects as regular sources of information and knowledge. For example, over a third of Internet users in the United States have consulted *Wikipedia*, and almost 10% consult it every day (Rainie and Tancer 2007). As a result, the epistemic status

Don Fallis

of such large collaborative projects is a pressing issue for the growing field of *applied epistemology* (cf. Goldman 1999, Bishop and Trout 2005, Fallis and Whitcomb forthcoming).<sup>2</sup> And the main question is whether large collaborative projects, such as *Wikipedia*, can be reliable sources of information.

Since *Wikipedia* lacks many of the editorial controls of traditional encyclopedias, many people (e.g., Keen 2007, Garfinkel 2008) have questioned the reliability of *Wikipedia*. For example, the former editor of *Encyclopædia Britannica*, Robert McHenry (2004) has famously denigrated it as the “faith-based encyclopedia.” And there certainly are legitimate reasons to worry about the reliability of *Wikipedia*. Unlike collaborations in science, Web 2.0 projects rarely restrict themselves to trained experts. These projects are typically open to anyone who is interested in participating. So, for example, there is no guarantee that the person writing or editing the *Wikipedia* article on bioethics has any training or expertise in bioethics. As a result, there is a distinct possibility that this contributor will introduce inaccurate information into the encyclopedia, or even remove accurate information from the encyclopedia (cf. Duguid 2006).<sup>3</sup> In addition, it is possible for people to use *Wikipedia* to engage in intentional deception (Seelye 2005). Finally, it is very easy for someone to simply delete an article in *Wikipedia* or replace it with gibberish or profanities.<sup>4</sup>

Despite these concerns, there is much theoretical and empirical evidence that large collaborative projects, such as *Wikipedia*, can actually be fairly reliable (cf. Surowiecki 2004, Sunstein 2006, Page 2007, Fallis 2008). When groups are sufficiently large and diverse, they can often come up with better information than the experts on a topic. For example, when a contestant on the television show *Who Wants to be a Millionaire?* is stumped by a question, she can poll the studio audience or phone a friend to get some help. It turns out that consulting the collective wisdom of the audience is a much more reliable “lifeline” than consulting your smartest friend (Surowiecki 2004, 4). And this phenomenon, often referred to as the *Wisdom of Crowds*, seems to apply to Web 2.0 projects.<sup>5</sup> For example, in a study sponsored by the journal *Nature* (Giles 2005) that involved a blind comparison by experts, the error rate for *Wikipedia* articles (on several scientific topics) was higher, but only *slightly* higher, than the error rate for Britannica articles.<sup>6</sup>

In any event, large collaborative projects that produce and disseminate information and knowledge are not going away any time soon.<sup>7</sup> Thus, it is critical to understand the *epistemology of mass collaboration*. Toward this end, the contributions to this issue of *Episteme* address the following important epistemological questions: How reliable are large collaborative projects that produce and disseminate information? What is the explanation for their reliability? Can large collaborative projects be reliable even if they do not make use of experts? Does the information produced by such projects count as *testimony*? Can we be justified in believing information produced by large collaborative projects? How should we go about deciding whether to believe information produced by such projects?

## PAPER SUMMARIES

As many philosophers (e.g., Lipton 1998, Lackey 2008) have pointed out, we acquire much of our knowledge from other people rather than from direct observation of the world. But the *epistemology of testimony* has tended to focus on one individual talking to another or possibly one individual talking to many others, in the case of books or television, for example. In “*Wikipedia* and the Epistemology of Testimony,” Deborah Tollefsen argues that groups in general, and *Wikipedia* in particular, can be sources of testimony. Tollefsen admits that many *Wikipedia* articles (e.g., those that are brand-new) may only count as the testimony of a single, anonymous individual or just a handful of such individuals. However, articles that have been read and edited by many contributors and that have reached a fairly stable state (e.g., the *Featured Articles*) have arguably been endorsed by the *Wikipedia* community as a whole. Tollefsen goes on to offer two possible accounts of how we can be justified in believing such testimony of the *Wikipedia* community.<sup>8</sup>

The next three contributions investigate why Web 2.0 projects, such as *Wikipedia*, have been successful, and whether they are successful *epistemically* (e.g., is *Wikipedia* a *reliable* source of testimony?). In “Web 2.0 vs. the Semantic Web: A Philosophical Assessment,” Luciano Floridi looks at the issue of why Web 2.0 in general has been successful. He does so by comparing Web 2.0 with another notable Internet project, *Web 3.0* or the *Semantic Web*. The goal of Web 3.0 is essentially to make the web intelligent (e.g., to automate the processing of semantic content). Floridi predicts that Web 3.0 will fail for the same reasons that grand artificial intelligence (AI) projects have failed in the past. However, this failure helps to explain why Web 2.0 has been, and probably will continue to be, a success. Instead of requiring the creation of artificial agents that can process semantic content, Web 2.0 works by making use of large numbers of agents that already can process semantic content (viz., human beings).

Modern science provides the most notable success story for collaboration in the production of knowledge. For example, in the physical sciences, almost all research is now collaborative. In fact, some research papers list hundreds of authors. As noted above, work in the philosophy of science has been done to explain why such collaborative scientific research has been successful at producing knowledge. In “The Epistemic Cultures of Science and *Wikipedia*: A Comparison,” K. Brad Wray argues that these explanations cannot be used to explain why *Wikipedia* will be successful at producing knowledge. This is largely because scientific researchers and *Wikipedia* contributors have very different goals, operate under very different social norms, and face very different incentive structures. For example, the goal of *Wikipedia* is to collect and disseminate existing knowledge rather than to discover new knowledge. And more importantly, *Wikipedia* contributors receive much less benefit than scientists for getting things right and suffer much less cost for getting things wrong. On the basis of such disanalogies, Wray concludes that we have good reason to worry about the reliability of *Wikipedia*. And even if *Wikipedia* is

Don Fallis

a reasonably reliable source of encyclopedic information, it is not reliable for the reasons that collaborative scientific research is reliable.

As noted above, many collaborative projects are successful despite the fact (and possibly because of the fact) that they are open to anyone who wants to participate. In “The Fate of Expertise after *Wikipedia*,” Lawrence Sanger (one of the founders of *Wikipedia*) argues that the success of *Wikipedia* does not show that society at large can do without experts. In fact, *Wikipedia* itself would not be nearly as reliable as it is without experts. For example, one of the reasons that *Wikipedia* is fairly reliable is that contributors are supposed to cite published sources for the claims that they make, and published sources typically make use of experts (as writers and editors).<sup>9</sup> Another reason that *Wikipedia* is fairly reliable is that many experts (even though they do not have an official role) do contribute. Finally, one of the reasons that *Wikipedia* is not as reliable as it might be is that experts are often deterred from participating in *Wikipedia*. Thus, Sanger suggests that we can potentially increase the reliability of such large collaborative projects by increasing the participation of experts. Sanger’s own Web 2.0 project, Citizendium.org, is intended to be an expert-friendly alternative to *Wikipedia*.<sup>10</sup>

While there is clearly a dispute about how reliable *Wikipedia* is, everyone has to admit that the quality of *Wikipedia* varies greatly from article to article. Some *Wikipedia* articles (e.g., the *Featured Articles*) are better than many published works. However, many articles are seriously incomplete, badly written, have no citations, etc. In “On Trusting *Wikipedia*,” P. D. Magnus considers how *Wikipedia* might be a useful source of information and knowledge despite this variability. *Wikipedia* clearly differs from traditional encyclopedias in several important respects (e.g., it does not have the same quality control mechanisms, it is more easily accessible, and it has much greater breadth). And Magnus describes how we put ourselves in epistemic danger if we use *Wikipedia* in the very same way that we use traditional encyclopedias. In particular, *Wikipedia* thwarts several standard techniques for verifying the accuracy of information. For example, when evaluating a source of information, we might consider the plausibility of its content or we might check its content against other independent sources. However, one of the first things that contributors to *Wikipedia* do when they edit an existing article is remove any claims that are clearly implausible. Also, since *Wikipedia* articles are so easily and so commonly copied by other sources on the web, it is not easy to find other sources that are clearly independent. Thus, Magnus suggests we need to develop new techniques for verifying the accuracy of information produced by large collaborative projects, such as *Wikipedia*.

Most collaborative projects settle on a collective view by building a consensus. Although the contributors do not meet face-to-face around a table at the same time, *Wikipedia* operates in essentially this way. A statement stands in *Wikipedia* once none of the contributors are interested in challenging it. In “Prediction Markets: The Practical and Normative Possibilities for the Social Production of Knowledge,” George Bragues considers an alternative mechanism that can

potentially make large collaborative projects even more reliable (at least with respect to predicting future events that can be precisely specified). Instead of building consensus, collaborative projects can settle on a collective view simply by taking a vote, or by aggregating the individual views in some similar fashion. This is essentially how Digg.com operates. And such an aggregation mechanism works even better when individuals have something at stake (e.g., money) when they express their views. For example, the Iowa Electronic Markets allow people to buy and sell contracts about future events (e.g., about whether a particular person will win a particular election). It turns out that the value of such a contract, which we can take to be the collective view of the market, is a very good indication of whether the event in question will actually occur (cf. Surowiecki 2004, 17–22).

## ACKNOWLEDGMENTS

I would like to thank Alvin Goldman, Kay Mathiesen, Larry Sanger, and all of the contributors for their invaluable assistance and advice in putting this issue together.

## REFERENCES

- Arico, Adam, Nathan Ballantyne, Matt Bedke, Jacob Caton, Ian Evans, Don Fallis, Brian Fiala, Martin Frické, David Glick, Peter Gross, Terry Horgan, Jenann Ismael, John Pollock, Daniel Sanderman, Paul Thorn, and Orlin Vakarelov.** 2008. "An Objectivist Argument for Thirdism." *Analysis* 68: 149–55.
- Bishop, Michael A. and J. D. Trout.** 2005. *Epistemology and the Psychology of Human Judgment*. Oxford: Oxford University Press.
- Cronin, Blaise, Debora Shaw, and Kathryn La Barre.** 2003. "A Cast of Thousands: Coauthorship and Subauthorship Collaboration in the 20th Century as Manifested in the Scholarly Journal Literature of Psychology and Philosophy." *Journal of the American Society for Information Science and Technology* 54: 855–71.
- Duguid, Paul.** 2006. "Limits of Self-Organization: Peer Production and 'Laws of Quality'." *First Monday*. Retrieved November 1, 2008 from [http://www.firstmonday.org/issues/issue11\\_10/duguid/](http://www.firstmonday.org/issues/issue11_10/duguid/)
- Encyclopædia Britannica.** 2006. "Fatally Flawed: Refuting the Recent Study on Encyclopedic Accuracy by the Journal *Nature*." Retrieved November 1, 2008 from [http://corporate.britannica.com/britannica\\_nature\\_response.pdf](http://corporate.britannica.com/britannica_nature_response.pdf)
- Fallis, Don.** 2006. "The Epistemic Costs and Benefits of Collaboration." *Southern Journal of Philosophy* 44(supplement): 197–208.
- Fallis, Don.** 2008. "Toward an Epistemology of *Wikipedia*." *Journal of the American Society for Information Science and Technology* 59: 1662–74.
- Fallis, Don and Dennis Whitcomb.** Forthcoming. "Epistemic Values and Information Management." *The Information Society*.
- Fischman, Josh.** 2008. "Encyclopædia Britannica Goes – Gasp! – Wiki." *Chronicle of Higher Education*. Retrieved November 1, 2008 from <http://chronicle.com/wiredcampus/article/3064/encyclopedia-britannica-goes-gasp-wiki>

Don Fallis

- Garfinkel, Simson L.** 2008. "Wikipedia and the Meaning of Truth." *Technology Review*. Retrieved November 1, 2008 from <http://www.technologyreview.com/web/21558/>
- Giles, Jim.** 2005. "Internet Encyclopaedias Go Head to Head." *Nature* 438: 900–1.
- Goldman, Alvin I.** 1999. *Knowledge in a Social World*. New York: Oxford University Press.
- Hardwig, John.** 1991. "The Role of Trust in Knowledge." *Journal of Philosophy* 88: 693–708.
- Keen, Andrew.** 2007. *Cult of the Amateur: How Today's Internet Is Killing Our Culture*. New York: Doubleday.
- Lackey, Jennifer.** 2008. *Learning From Words*. Oxford: Oxford University Press.
- Lipton, Peter.** 1998. "The Epistemology of Testimony." *Studies in the History and Philosophy of Science* 29: 1–31.
- McHenry, Robert.** 2004. "The Faith-Based Encyclopedia." *TCS Daily*. Retrieved November 1, 2008 from <http://www.tcsdaily.com/article.aspx?id=111504A>
- Page, Scott E.** 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton: Princeton University Press.
- Rainie, Lee and Bill Tancer.** 2007. "Data Memo." *PEW Internet & American Life Project*. Retrieved November 1, 2008 from [http://www.pewinternet.org/pdfs/PIP\\_Wikipedia07.pdf](http://www.pewinternet.org/pdfs/PIP_Wikipedia07.pdf)
- Schauer, Frederick.** 2002. "The Dilemma of Ignorance: PGA Tour, Inc. v. Casey Martin." *Supreme Court Review* 2001: 267–97.
- Seelye, Katharine Q.** 2005. "A Little Sleuthing Unmasks Writer of Wikipedia Prank." *New York Times*. Retrieved November 1, 2008 from <http://www.nytimes.com/2005/12/11/business/media/11web.html>
- Sunstein, Cass R.** 2006. *Infotopia: How Many Minds Produce Knowledge*. Oxford: Oxford University Press.
- Surowiecki, James.** 2004. *The Wisdom of Crowds*. New York: Doubleday.
- Thagard, Paul.** 1997. "Collaborative Knowledge." *Nous* 31: 242–61.
- Viñas, Maria J.** 2008. "Medical Wiki Backed by Prominent Colleges Will Go Live by Year's End." *Chronicle of Higher Education*. Retrieved November 1, 2008 from <http://chronicle.com/free/v35/i02/02a01702.htm>
- Wray, K. B.** 2002. "The Epistemic Significance of Collaborative Research." *Philosophy of Science* 69: 150–68.

---

## NOTES

- 1 Collaborative research is increasing in the sciences and in many other disciplines. There is even some formal collaboration in philosophy (e.g., Arico et al. 2008). There is not as much here as there is in the sciences (cf. Fallis 2006). However, even in philosophy, there is quite a bit *informal* collaboration (cf. Cronin et al. 2003).
- 2 Just as applied ethics addresses concrete, practical issues from a moral standpoint, applied epistemology addresses concrete, practical issues from an epistemological standpoint.
- 3 Simson Garfinkel (2008) has criticized *Wikipedia* by pointing to a case where the subject of an article was not able to correct a statement about himself that he very well knew to be false. His attempts to correct the inaccuracy would quickly be reversed because he could not cite a published source that supported his position. Although this may sound

## INTRODUCTION

problematic on the face of it, it is not clear that this is really a serious problem or that it is unique to *Wikipedia*. Most encyclopedias stick to published sources, and are arguably more reliable for doing so. But published sources will not always have the most current or most accurate information in particular cases.

- 4 One of the great advantages of the technology used in *Wikipedia* is that it is very easy for legitimate contributors to find and correct such vandalism.
- 5 However, it is not clear that *Wikipedia* has all of the characteristics that Surowiecki identifies as being distinguishing features of wise crowds (Fallis 2008, 1670).
- 6 Encyclopædia Britannica (2006) has criticized this study on a number of points.
- 7 In fact, even Encyclopædia Britannica, which has long been critical of *Wikipedia*, is planning to allow users to contribute and edit articles that will be available alongside the standard edited content of the encyclopedia (Fischman 2008).
- 8 Some Web 2.0 projects do not really require us to rely on group testimony. For example, while it is fairly difficult to find large primes (and, thus, it is useful to have large collaborative efforts to find them), it is relatively easy to check whether a number that is purported to be prime really is. However, it is rather more difficult for users to directly confirm the information provided by *Wikipedia* or Yahoo! Answers.
- 9 Admittedly, citing published sources does not guarantee reliability. For example, in its decision on *Bush v. Gore*, the United States Supreme Court relied on election statistics from an article in an Omaha newspaper written by a young reporter who was so junior that his office was a closet (Schauer 2002, 287). While these statistics may have been accurate, it is probably not a good idea to rely on such a source when one is making such a critical decision.
- 10 A collaborative online medical encyclopedia, Medpedia.com, is in the works which will *only* allow experts to contribute (Viñas 2008). However, this strategy may have epistemic costs that outweigh its epistemic benefits in terms of greater reliability (Fallis 2008, 1671).

Don Fallis is Associate Professor of Information Resources and Adjunct Associate Professor of Philosophy at the University of Arizona. His articles on epistemology have appeared in the *Journal of Philosophy*, *Philosophical Studies*, *Episteme*, and the *Journal of the American Society for Information Science and Technology*.