Toward an Epistemology of *Wikipedia*

Don Fallis

School of Information Resources and Library Science

University of Arizona

This is a preprint of an article published in the

Journal of the American Society for Information Science and Technology, **59**, 10, 1662-74

http://www.interscience.Wiley.com/

Toward an Epistemology of *Wikipedia*

Abstract: *Wikipedia* (the "free online encyclopedia that anyone can edit") is having a huge impact on how a great many people gather information about the world. So, it is important for epistemologists and information scientists to ask whether or not people are likely to acquire knowledge as a result of having access to this information source. In other words, is Wikipedia having good *epistemic* consequences? After surveying the various concerns that have been raised about the reliability of Wikipedia, this paper argues that the epistemic consequences of people using Wikipedia as a source of information are likely to be quite good. According to several empirical studies, the reliability of Wikipedia compares favorably to the reliability of traditional encyclopedias. Furthermore, the reliability of Wikipedia compares even more favorably to the reliability of those information sources that people would be likely to use if Wikipedia did not exist (viz., websites that are as freely and easily accessible as Wikipedia). In addition, Wikipedia has a number of other epistemic virtues (e.g., power, speed, and fecundity) that arguably outweigh any deficiency in terms of reliability. Even so, epistemologists and information scientists should certainly be trying to identify changes (or alternatives) to Wikipedia that will bring about even better epistemic consequences. This paper suggests that, in order to improve Wikipedia, we need to clarify what our epistemic values are and we need a better understanding of why Wikipedia works as well as it does.

> *Somebody who reads Wikipedia is "rather in the position of a visitor to a public restroom," says Mr. McHenry, Britannica's former editor. "It may be obviously dirty, so that he knows to exercise great care, or it may seem fairly clean, so that he may be lulled into a false sense of security. What he certainly does not know is who has used the facilities before*

> *him." One wonders whether people like Mr. McHenry would prefer there to be no public lavatories at all.*
>
> – *The Economist* (April 22, 2006)

1. Introduction

Mass collaboration is one of the newest trends in the creation and dissemination of knowledge and information. Several people working together to produce knowledge is nothing new, of course (cf. Thagard 2006). But until recently, such projects have been limited in the number of collaborators that can participate and in the distance between them. It is now possible for millions of people separated by thousands of miles to collaborate on a single project.[1] *Wikis*, which are websites that anyone with Internet access can edit, provide a popular medium for this sort of collaboration.

Such mass collaboration has been extremely successful in many instances (cf. Tapscott and Williams 2006). A popular example is the development of open source software, such as the *Linux* operating system (cf. Duguid 2006). Another nice example is the *Great Internet Mersenne Prime Search* (see http://www.mersenne.org). By allowing anybody with an *Intel Pentium* processor to participate in the search, this project has discovered several of the largest known prime numbers.

However, it is not a foregone conclusion that such mass collaboration will be successful in all instances. For example, it seems unlikely that a million people working together would write a very good novel. But are a million people working together likely to compile a good *encyclopedia*? This paper investigates the success of a notable example of mass collaboration on the Internet: the "free online encyclopedia that anyone can edit" *Wikipedia*. After discussing the various concerns that have been raised about the quality of the information on Wikipedia, I argue that people are likely to acquire knowledge as a result of having access to this information source.

2. An Epistemic Evaluation of Wikipedia

There are actually a number of different ways in which a project like Wikipedia might (or might not) be successful. This project might be successful at building a good encyclopedia. But it might also be successful at simply building an online community. And it is not completely clear which of these goals has priority (cf. Sanger 2005). In fact, one of the founders of Wikipedia, Jimmy Wales, has said that "the goal of Wikipedia is *fun* for the contributors" (quoted in Poe 2006, emphasis added). Even if the contributors to Wikipedia ultimately just want to have fun, however, building a good encyclopedia is still an important goal of this project. Similarly, even if the owners of Encyclopedia Britannica ultimately just want to make money, building a good encyclopedia is still an important goal that they have. And this goal is clearly *epistemic*. A good encyclopedia is a place where people can "acquire knowledge" and sometimes "share knowledge" (http://en.wikipedia.org/wiki/Wikipedia:About).[2]

Epistemology is the study of what knowledge is and how people can acquire it (cf. Feldman 2003). And, according to Alvin Goldman (1999), a primary task for the epistemologist is to evaluate institutions, such as Wikipedia, in terms of their epistemic consequences. That is, the epistemologist should ask whether people are more (or less) likely to acquire knowledge as a result of a particular institution being in existence. Several people (e.g., Shera 1970, 82-110, Wilson 1983, Goldman 1999, 161-188, Fallis 2006) have pointed out that such work in epistemology can be critical to information science. In particular, it is important to know whether people are likely to acquire knowledge from popular information sources. For example, library and information scientists frequently evaluate reference service in terms of its reliability (cf. Meola 2000). Also, Goldman (2008) has recently looked at the epistemic consequences of blogging (as compared to the conventional news media). In a similar vein, this paper investigates the epistemic consequences of Wikipedia.

Wikipedia certainly has the potential to have great epistemic benefits. By allowing anyone with Internet access to create and edit content (i.e., by taking advantage of what Chris Anderson (2006, 219) calls "crowdsourcing"), Wikipedia now includes millions of entries in

many different languages.[3] Because all of this material is freely and easily accessible by anyone with Internet access, Wikipedia is now one of the top ten Internet domains in terms of Internet traffic along with *Google*, *Yahoo*, *YouTube*, and *MySpace*. Over a third of Internet users in the United States have consulted Wikipedia, and almost 10% consult it every day (cf. Rainie and Tancer 2007). It essentially serves as an aggregation point for encyclopedic information in much the same way that the online auction website *Ebay.com* serves as an aggregation point for other goods (cf. Anderson 2006, 89).

Nevertheless, serious concerns have been raised about the quality (accuracy, completeness, comprehensibility, etc.) of the information on Wikipedia. Entries in traditional encyclopedias are often written by people with expertise on the topic in question. Also, these entries are checked for accuracy by experienced editors before they are published. However, because it allows anyone with Internet access to create and modify content, Wikipedia lacks these sorts of quality control mechanisms. In fact, "no one stands officially behind the authenticity and accuracy of any information in [Wikipedia]" (Denning et al. 2005). As a result, it has been suggested that "it's the blind leading the blind—infinite monkeys providing infinite information for infinite readers, perpetuating the cycle of misinformation and ignorance" (Keen 2007, 4). Wikipedia has also been dismissed as unreliable by members of the library and information science community (e.g., Cronin 2005, Gorman 2007). The ultimate worry here is that people are likely to acquire false beliefs rather than knowledge as a result of consulting such a seemingly unreliable information source.

3. The Epistemology of Encyclopedias

A good encyclopedia should help people to acquire knowledge. Thus, encyclopedias are clearly of interest to the epistemologist. Before we look at how well Wikipedia in particular helps people to acquire knowledge, it will be useful to say more precisely how encyclopedias in general might be studied by an epistemologist. As I describe in this section, it turns out that the

epistemology of encyclopedias must appeal to work on several important topics in current epistemological research: viz., the *epistemology of testimony*, *social epistemology*, and *epistemic value theory*. But it should be noted that, since epistemology has not really addressed the issue of mass collaboration, an epistemic evaluation of Wikipedia will require an extension, and not just an application, of the existing epistemological research.

Epistemology has traditionally been fairly *individualistic*. That is, it typically focuses on how an individual working alone can acquire knowledge (cf. Goldman 1999, 4). Yet we acquire most of our knowledge from other people rather than from direct observation of the world (cf. Hume 1977 [1748], 74, Wilson 1983). And encyclopedias are in the business of transmitting information and knowledge from one group of people to another group of people. In other words, their goal is to disseminate existing knowledge rather than to discover new knowledge.[4] Thus, the epistemology of encyclopedias falls within the scope of the *epistemology of testimony*. The epistemology of testimony looks at how it is possible to come to know something based solely on the fact that somebody else says that it is so (cf. Lackey and Sosa 2006).

Whereas the epistemology of testimony typically focuses on a single individual transmitting knowledge to another individual, the epistemology of encyclopedias is clearly much more social than that. With encyclopedias (as well as most other types of recorded information such as books and newspapers), the receiver of the information is not a single individual. As a result, we have to be concerned with things like *how many* people are able to acquire knowledge from such sources (cf. Goldman 1999, 93-94). In addition, the source of the information in an encyclopedia is rarely a single individual. First, an encyclopedia entry is almost never the original source of the information that it contains. Encyclopedias collect, condense, and organize knowledge that many other people have already discovered. In fact, contributors to Wikipedia are prohibited from including any original results (see http://en.wikipedia.org/wiki/Wikipedia:No_original_research). Second, encyclopedias are rarely created by a single individual working alone. Instead, they are an example of *group testimony*

(cf. Tollefsen 2007). Many authors, editors, artists, and designers are needed to create a modern encyclopedia (cf. Pang 1998). And Wikipedia simply extends this even further by allowing anyone with Internet access to participate in the project. Thus, the epistemology of encyclopedias also falls within the scope of *social epistemology*. Most work in epistemology is primarily concerned with how cognitive and perceptual processes *within an individual* lead to the acquisition of knowledge. In contrast, social epistemology looks at how *social* processes lead to the acquisition of knowledge (cf. Shera 1970, 82-110, Goldman 1999).

In point of fact, the creation and dissemination of knowledge in general is very often a social activity. For example, libraries, publishing companies, universities, search engine companies, and even *dictionaries* are typically large-scale collective endeavors. While early lexicographers, such as Samuel Johnson, worked largely on their own, subsequent dictionaries have always been produced by large teams. In fact, in its early days, the *Oxford English Dictionary*, in a strategy very similar to Wikipedia, solicited help from the general public (cf. Winchester 1998, 101-114).

Although interest in this topic goes back to Aristotle (cf. Waldron 1995), it should be noted that there is actually not a lot of philosophical work on how people come together to collaboratively create information and knowledge. Some work (e.g., in the philosophy of science) has been done on small-scale collaboration (cf. Wray 2002, Thagard 2006). Also, there is work on simply *aggregating* (e.g., by averaging or by taking a majority vote) the views of many individuals that I will appeal to below (cf. Estlund 1994, Goldman 1999, 81-82, Surowiecki 2004). In addition, some work has been done in information science on *collaboration* and on *communal labor* (cf. Warner 2005, 557, Birnholtz 2007). But only a very few social epistemologists (e.g., Magnus 2006, the present paper, and a future issue of the journal *Episteme*) have begun to address the sort of *mass collaboration* on a single project that takes place in Wikipedia.

Finally, if we want to evaluate any social institution in terms of its epistemic consequences, we also need to know what counts as *good* epistemic consequences (cf. Goldman 1999, 87-100). In other words, we need to know what sorts of things are epistemically valuable. Thus, the epistemology of encyclopedias must appeal to work on *epistemic values* (cf. Riggs 2003, Pritchard 2007).

Philosophical research on epistemic values tends to focus on perennial questions in epistemology that go back to Plato (1961). For example, exactly why is *knowing* something more valuable than simply having a *true belief* about it? But this research has rarely been applied to actual decisions that people make where epistemic consequences are at stake. For example, when someone purchases an encyclopedia, is she more concerned with how much accurate information it contains or with how much of its information is accurate? There is some philosophical work on the "epistemic utilities" of scientists (cf. Levi 1967, Maher 1993). But only a few people have looked at what epistemic values are at play in more mundane contexts, such as education and information management (cf. Paterson 1979, Goldman 1999, Fallis 2004a, Fallis 2006, 495-503). As a result, it is difficult to say precisely what counts as a good epistemic consequence in these contexts. As I endeavor to show below, however, enough is now known about epistemic values to argue convincingly that the epistemic benefits of Wikipedia outweigh its epistemic costs. I begin by laying out why so many people have thought that Wikipedia will not have good epistemic consequences.

4. Epistemic Concerns about Wikipedia

There are several dimensions of information quality: *accuracy*, *completeness*, *currency*, *comprehensibility*, etc. (cf. Fox 1994). It has been suggested that the information on Wikipedia fails on several of these dimensions. For example, Wikipedia entries are sometimes badly written and important topics are not always covered (cf. Rosenzweig 2006).[5] In other words, Wikipedia

is not as comprehensible and complete as we might expect an encyclopedia to be. It is clear that such failings can adversely affect people's ability to acquire knowledge from Wikipedia.

However, *inaccurate* information can easily lead people to acquire false beliefs. In other words, inaccurate information can make people epistemically worse off instead of just failing to make them epistemically better off. And epistemologists (e.g., Hume 1977 [1748], 111, Descartes 1996 [1641], 12) typically consider falling into *error* to be the most adverse epistemic consequence (cf. Riggs 2003, 347). Thus, the principle epistemic concern that has been raised about Wikipedia is whether people are likely to get *accurate* information from it. In other words, is Wikipedia a reliable source of information? (An information source is *reliable* if most of the information that it contains is accurate.)

As noted above, more and more people are using Wikipedia as a source of information. It has even been cited in court cases (cf. Cohen 2007). But concerns about its reliability in particular have led many people (e.g., Denning et al. 2005, Keen 2007) to suggest that Wikipedia should not be used as a source of information. In fact, the history department at Middlebury College has forbidden its students from citing Wikipedia (cf. Read 2007a). In this section, I describe the various reasons to worry about the reliability of Wikipedia (and about the related issue of its verifiability).

It should be noted that library and information scientists (e.g., Wilson 1983, 21-26, Rieh 2002, Kelton et al. 2008) often focus on the *descriptive* question of when people actually do grant *cognitive authority* to information sources. We are concerned here with the *normative* question, from the epistemology of testimony, of whether people *ought* to grant cognitive authority to Wikipedia (cf. Goldman 2001, Fallis 2004b, 468). In other words, instead of looking at the conditions under which people trust information sources, we want to know with whether this particular information source really is *trustworthy*.

a) Concerns about its Reliability

Wikipedia differs from many other collaborative projects in that it does not directly bump up against reality (cf. Duguid 2006). For example, in order for software to be added to the *Linux* operating system, it actually has to work. Similarly, if a participant in the *Great Internet Mersenne Prime Search* claims to have discovered the largest known prime number, it is a simple matter to check that the number really is prime before announcing it to the world. By contrast, information can be added to Wikipedia and remain on Wikipedia indefinitely regardless of whether or not it is accurate.

There are several reasons to think that a significant amount of information on Wikipedia might be inaccurate. First, since anyone can contribute to Wikipedia, many of these contributors will not have much expertise in the topics that they write about. Also, because people can contribute anonymously, some of those who claim to have expertise or credentials do not (cf. Schiff 2006, Read 2007b). Given their lack of expertise, such contributors may inadvertently add inaccurate information to Wikipedia. In addition, they may inadvertently remove accurate information (cf. Duguid 2006). Thus, there may be some amount of *misinformation* on Wikipedia.

Moreover, the problem is not just that Wikipedia allows people who lack expertise to contribute. It has been suggested that Wikipedia exhibits *anti-intellectualism* and actively deters people with expertise from contributing. For example, experts rarely receive any deference from other contributors to Wikipedia as a result of their expertise (cf. Keen 2007, 43). Since they cannot simply appeal to their authority, experts have to fight it out just like anyone else to get their views to stick in the encyclopedia. Many experts are understandably unwilling to put in the effort to create content that might simply be removed by an unqualified individual with an axe to grind. Furthermore, academics and other experts who create information and knowledge typically want to get credit for their work (cf. Goldman 1999, 260). But since Wikipedia entries are the creation of multiple (often anonymous) authors and editors, no one person can claim credit for the result.

Second, since anyone can contribute to Wikipedia, some of these contributors may try to deceive the readers of Wikipedia.[6] For example, the entry on the journalist John Siegenthaler was famously modified to falsely claim that he was involved in the Kennedy assassinations (cf. Sunstein 2006, 156). And this inaccurate information was on the website for over four months. As Peter Hernon (1995, 134) puts it, "inaccurate information might result from either a deliberate attempt to deceive or mislead (*disinformation*), or an honest mistake (*misinformation*)." Thus, there may also be some amount of *disinformation* on Wikipedia.

Whenever someone has an interest in convincing other people to believe something even if it is not true, there is some reason to worry about the accuracy of the information that she provides (cf. Fallis 2004b, 469). For example, it is worrying when prominent individuals (e.g., members of Congress) and large organizations are caught changing their own Wikipedia entries (cf. Schiff 2006, Borland 2007, Keen 2007, 4). And even if someone is not engaged in outright deception, there is still potential for inaccurate information to be introduced as a result of unintentional bias (cf. Goldman 2001, 104-105).

Finally, there is a third category of inaccurate information that may be found on Wikipedia. Since we can all edit Wikipedia, Stephen Colbert (host of the satirical television news show *The Colbert Report*) has suggested that we should just construct the reality that we collectively want (see http://colbertondemand.com/videos/The_Word/The_Word_Wikiality). For example, since we are all concerned with the survival of endangered species, Colbert encouraged his viewers to edit the entry on African Elephants to say that their numbers had tripled in the last six months. This type of inaccurate information is arguably distinct from both disinformation and misinformation. Unlike someone who intends to deceive or who makes an honest mistake, Colbert shows no concern for the *truth* of this Wikipedia entry. (Someone who intends to deceive is concerned to avoid the truth.) Several philosophers (e.g., Black 1983, Cohen 2002, Frankfurt 2005) have offered analyses of *humbug* or *bullshit*. And, according to Harry Frankfurt (2005, 33-34), it is "this lack of connection to a concern with truth—this indifference to how things really

are—that I regard as of the essence of bullshit." Thus, there may be some amount of bullshit on Wikipedia.[7]

b) Concerns about its Verifiability

The main reason that the reliability of Wikipedia is a concern is that people can be misled by inaccurate information. And being misled can often lead to serious harms (cf. Fallis 2004b, 465). But inaccurate information is not so serious a problem if it is possible for people to determine that this information is (or is very likely to be) inaccurate. In other words, if people are in a position to verify the accuracy of information, they are less likely to be misled by inaccurate information. Hence, we need to consider the verifiability as well as the reliability of this information source. (An information source is *verifiable* if people can easily determine whether the information that it contains is accurate (cf. Fallis 2004b, 476-477).)

It is important to note that people can avoid the potential epistemic costs of inaccurate information even if they are not able to determine with absolute certainty that a particular piece of information is inaccurate. It is often sufficient for people to have a reasonable estimate of the reliability of the source of the information. There can certainly be bad epistemic consequences if our estimate of the reliability of a source does not match its actual reliability. For example, relevant evidence is often withheld from juries because they are likely to overestimate the probative value of this evidence (cf. Goldman 1999, 294-295). But if we have the right amount of faith in them, even fairly unreliable sources can be useful (cf. Goldman 1999, 121, Fallis 2004b, 473-474). For example, we might simply raise our degree of confidence in claims made by such sources without fully accepting that these claims are true.

P. D. Magnus (2006), however, has raised concerns about the verifiability of Wikipedia. He points out that we can try to verify the accuracy of a particular claim (that we are uncertain about) by considering both the *presentation* and the *content* of the information (cf. Fallis 2004b, 472-473). For instance, if an author makes numerous spelling and grammatical mistakes or

makes other claims that are clearly false, then we have reason to be cautious about the accuracy of this particular claim. Unfortunately, these are just the sorts of features that contributors to Wikipedia typically remove when they edit entries that other people have written. That is, they quickly remove spelling mistakes, grammatical mistakes, and clearly implausible claims. Thus, these features will no longer be available to someone trying to verify the accuracy of these entries. To use Mr. McHenry's analogy, the concern is that people cannot tell how dirty a restroom really is because others have come through ahead of them, picked up the trash, and quickly wiped off the counters.

We can also try to verify the accuracy of a piece of information by considering the identity of the source of that information (cf. Fallis 2004b, 469-470). Does this source have any *conflict of interest* that might lead her to intentionally disseminate inaccurate information on this topic? Also, is this source sufficiently *qualified* (or have a good enough *track record*) on this topic that she would be unlikely to unintentionally disseminate inaccurate information? But, in the case of Wikipedia, it is rather difficult to determine exactly who the source of a particular piece of information is. Any given entry may have been edited by several different contributors and Wikipedia allows these contributors to remain anonymous if they wish. (Contributors do not have to register and, even if they do, they can simply pick a user name that hides their identity.)

5. Wikipedia is quite Reliable

Despite legitimate concerns about its reliability, empirical evidence actually suggests that Wikipedia is not all that unreliable. For instance, researchers have tested Wikipedia by inserting plausible errors and seeing how long it takes for the errors to be corrected (cf. Read 2006). While there are exceptional cases (see http://www.frozennorth.org/C2011481421/E652809545/), such vandalism is typically corrected in just a few minutes (cf. Viégas et al. 2004, 579). In addition, blind comparisons by experts of Wikipedia entries and entries in a traditional encyclopedia have

been carried out. For example, a study (Giles 2005) by the journal *Nature* found that Wikipedia was only slightly less reliable than *Encyclopedia Britannica*.[8]

To be fair, it should be noted that the *Nature* study focused specifically on entries on scientific topics. Results have been more mixed when it comes to other topics. For example, in a blind comparison of Wikipedia and Britannica with respect to a small selection of entries on *philosophical* topics, Magnus (2006) found that "Wikipedia entries vary widely in quality." George Bragues (2007), who evaluated the Wikipedia entries on seven great philosophers using authoritative reference works on these philosophers, reached the same conclusion. Nevertheless, even on such non-scientific topics, the *reliability* of Wikipedia still seems to be comparable to that of Britannica. For example, Magnus found that, while Wikipedia had more "major errors," Britannica had many more "minor errors and infelicities." And, in fact, Bragues was "unable to uncover any outright errors [in Wikipedia]. The sins of Wikipedia are more of omission than commission." Similarly, a study (Devgan et al. 2007) that looked at Wikipedia entries on medical topics found no inaccuracies, but found several significant omissions.

Several investigations (e.g., Rosenzweig 2006, Bragues 2007) have simply rated the quality of the information on Wikipedia (by consulting authorities or authoritative sources). However, it is often more appropriate to carry out a *relative* rather than an *absolute* epistemic evaluation of an institution (cf. Goldman 1999, 92-93). In other words, rather than simply determining exactly how reliable an information source is, we should really determine how reliable it is compared to the available alternatives. Therefore, it makes sense to compare the reliability of Wikipedia to the reliability of traditional encyclopedias as *Nature* and Magnus have done.

As Bertrand Meyer (2006) has suggested, however, it is not clear that this is the most appropriate comparison to make. He points out that we should really be comparing the reliability of Wikipedia against the reliability of the information sources that people would likely be using if Wikipedia were not available: viz., the freely available websites on their topic of interest returned

by their favorite search engine. It is this comparison that will tell us whether it is, as a matter of fact, epistemically better for people to have access to Wikipedia. And, if the reliability of Wikipedia is comparable to the reliability of traditional encyclopedias, then the reliability of Wikipedia presumably compares even more favorably to the reliability of randomly chosen websites.[9] Several empirical studies (e.g., Impicciatore et al. 1997, Fallis and Frické 2002) have found significant amounts of inaccurate information on the Internet. And websites in general are not checked as quickly (or by as many people) as is Wikipedia.

Finally, it is important to note that the degree of reliability that we demand of an information source often depends on the circumstances (cf. Fallis 2004a, 111). For example, when we are seeking information out of pure curiosity, it may not be a big deal if some of this information turns out to be inaccurate. But when we are seeking information in order to decide on a medical treatment or a large investment, we would like to be sure that the information is accurate. And if reliability is sufficiently important, we should probably double check the information (e.g., by consulting an independent source of information) that we get from any encyclopedia. It is often suggested that encyclopedias should be a starting point rather than an ending point for research (cf. Anderson 2006, 69).

People will not always double check information even when the stakes are reasonably high. In other words, people are subject to the so-called *Principle of Least Effort*. Empirical studies have found that "most researchers (even "serious" scholars) will tend to choose easily available information sources, even when they are objectively of low quality" (Mann 1993, 91). As a result, the easy availability of low quality information sources can certainly have bad epistemic consequences in actual practice. But it is also important to emphasize that people do not just have *epistemic* interests. And given that people have many non-epistemic interests (e.g., they want to save time and money), it may often be rational not to seek more knowledge or greater justification (cf. Hardin 2003). Consequently, Wikipedia may be sufficiently reliable for many purposes even if it is not quite as reliable as a traditional encyclopedia.

6. Wikipedia is quite Verifiable

In this section, I argue that Wikipedia is also quite verifiable. While people may not always verify the accuracy of information on Wikipedia when they ought to, it is not clear that Magnus is correct that the tools that they need to do so are not available. First of all, empirical studies (e.g., Fallis and Frické 2002) indicate that spelling and grammatical mistakes are not correlated with inaccuracy. So, when such mistakes are removed from a Wikipedia entry, it is not clear that people have been deprived of useful indicators of inaccuracy. In addition, many other infelicities of presentation (e.g., lack of stylistic sophistication and coherence) are not as easily removed by other contributors. With regard to the removal of implausible claims, some people probably are being deprived of useful indicators of inaccuracy. Even so, claims that are clearly implausible to one person may not be clearly implausible to another. Hence, we have to weigh the epistemic cost of a loss of verifiability for some people against the epistemic benefit of removing information that will be misleading to other people.

It is certainly not easy to determine the real-life identity of the author of a specific Wikipedia entry. But it is not clear that this seriously impedes our ability to verify the accuracy of the entry. First, it should be noted that it is also not very easy to determine the real-life identity of the author of a specific entry in a traditional encyclopedia. With the exception of those encyclopedias that focus on a particular subject area, traditional encyclopedias rarely list the authors of each entry. Second, unlike a traditional encyclopedia, readers of Wikipedia can easily look at all of the contributions that a particular author has made and can evaluate the quality of these contributions. In any event, even if we could easily determine the real-life identity of an author, it would still be much too time consuming to research her qualifications and potential biases. We typically trust a particular encyclopedia entry, not because we trust its author, but because we trust the process by which the entries in the encyclopedia are produced. More generally, we trust *group testimony* if we know that the process by which the testimony is

produced is reliable (cf. Tollefsen 2007, 306-308). And the process by which entries in Wikipedia are produced seems to be fairly reliable.

Admittedly, the process may not be as reliable as the process used by traditional encyclopedias. But Wikipedia warns readers about the fact that it may contain inaccurate information (see http://en.wikipedia.org/wiki/Wikipedia:Disclaimers). (Traditional encyclopedias often do have similar disclaimers, but they are rarely displayed as prominently as those in Wikipedia.) And most people seem to be aware of this fact. By contrast, traditional encyclopedias (e.g., Encyclopedia Britannica 2006) often insist on their high level of accuracy. But the empirical studies discussed above suggest that there are many errors in traditional encyclopedias as well as in Wikipedia. As a result, there is reason to think that people are more likely to overestimate the reliability of traditional encyclopedias than the reliability of Wikipedia. As Eli Guinnee (2007) puts it, "an inaccuracy in Britannica is (mis)taken as fact, an inaccuracy in Wikipedia is taken with a grain of salt, easily confirmed or proved wrong."

As noted above, another way that we can try to verify the accuracy of a piece of information is by checking to see if other information sources corroborate the claim (cf. Fallis 2004b, 470-471). In the case of Wikipedia, this can be tricky because many other information sources on the Internet (such as *About.com* and *Answers.com*) have simply copied much of their content from Wikipedia (cf. Magnus 2006). Such websites are examples of what Goldman (2001, 102) would call "non-discriminating reflectors" of Wikipedia. But it is fairly easy for people to recognize that a website containing a word-for-word copy of a Wikipedia entry is not really an independent source and that it does not provide much corroboration.[10]

Finally, in many respects, Wikipedia is actually more verifiable than most other information sources. For example, in addition to general disclaimers, *Warnings* are placed at the top of Wikipedia entries whose accuracy or neutrality has been disputed. Also, unlike traditional encyclopedias, Wikipedia is not a black box. Readers of Wikipedia have easy access to the entire editing history of every entry. In addition, readers have access to the *talk pages* that contributors

use to discuss how entries should be changed (cf. Sunstein 2006, 152). Admittedly, most readers are only going to consult the current entry itself. But if someone is particularly interested in a topic, the editing history and the talk pages can be invaluable resources. For example, one can look to see if there were any dissenting opinions, what these different viewpoints were, and what arguments ultimately carried the day. This is right in line with John Stuart Mill's (1978 [1859]) claim that exposure to different viewpoints is the best way to learn the truth about a topic (cf. Goldman 1999, 212-213).

New technologies are also being developed that have the potential to increase the verifiability of Wikipedia (cf. Giles 2007).[11] In particular, software tools are being developed that allow readers to better estimate the reliability of authors of particular entries. For example, B. Thomas Adler and Luca de Alfaro (2007) have developed a system that automatically keeps track of each author's reputation for accuracy. Basically, authors gain in reputation if their contributions to Wikipedia persist and authors lose in reputation if their contributions to Wikipedia are quickly removed by other contributors.

In addition, Anthony et al. (2007) found that registered contributors who have made a lot of changes are likely to make changes that persist. Interestingly, anonymous contributors who have made only a few changes are also likely to make changes that persist. Since it is easy to determine which category a particular contributor falls into, these sorts of empirical results might also help increase the verifiability of Wikipedia.

Virgil Griffith has also created a searchable database (*Wikiscanner*) that allows readers to connect specific contributions to Wikipedia with the organization that owns the IP addresses from which those contributions originated (cf. Borland 2007). So, for example, readers can easily find out if employees of the *Diebold Corporation* have been editing the Wikipedia entry on the Diebold Corporation. In fact, Wikiscanner may also increase the reliability as well as the verifiability of Wikipedia by deterring people from editing entries when they have an obvious conflict of interest.

Still, it should be noted that there are some epistemic benefits to anonymity that we might lose as a result of tools like Wikiscanner. In particular, if people cannot disseminate information anonymously, they may be deterred from disseminating valuable information. For example, suppose that I have something critical to say about a person or an organization. If the target of my criticism can find out that I am the source (and might come after me), I may very well decide that I am better off not saying anything (cf. Kronick 1988, 225).

7. Wikipedia has many other Epistemic Virtues

Concerns about Wikipedia usually focus on its reliability (or lack thereof). But there are many other epistemic virtues beyond reliability. For example, in addition to reliability, Goldman has discussed the epistemic values of *power*, *speed*, and *fecundity* (cf. Thagard 1997). That is, we are also concerned with *how much* knowledge can be acquired from an information source, *how fast* that knowledge can be acquired, and *how many* people can acquire that knowledge. Fecundity, in particular, is an especially important concept at the intersection of social epistemology and work on epistemic values. As noted above, this epistemic value is critical to the epistemology of encyclopedias.

Wikipedia seems to do pretty well with regard to these other epistemic values (cf. Cross 2006, Guinnee 2007). Because it has a huge amount of free labor working around the clock, it is likely to be very powerful. Wikipedia is, for example, several times larger than Encyclopedia Britannica (cf. Cross 2006). Because there is no delay for new content to go through an editorial filter and because its content can be accessed quickly over the Internet, it is likely to be very speedy (cf. Thagard 1997, Sunstein 2006, 150). And because it is free to anyone with Internet access, it is likely to be very fecund.

In addition to being freely accessible, Wikipedia is essentially an "open source" project. That is, anyone is allowed to make copies of its content for free (and many websites do). While this increases the dissemination of knowledge, contributors do have to give up any intellectual

property rights that they might have claimed in their work. Intellectual property rights can have epistemic benefits by motivating people to create new content (cf. Fallis 2007, 36). But a large number of people are willing to contribute to Wikipedia even without this motivation (cf. Sunstein 2006, 153-154). In fact, if someone is more interested in disseminating her knowledge as widely as possible than in making money, Wikipedia will be a very attractive outlet (cf. Sanger 2005).

In addition, it is important to note that intellectual property rights can stifle as well as promote the creation of new content (cf. Fallis 2007, 37). For example, since a copyright holder has the exclusive right to produce derivative works, other people are deterred from engaging in such creative efforts. Since its content can be freely copied and modified, Wikipedia avoids such epistemic costs. In fact, contributors to Wikipedia are encouraged to edit the content that other people have created.

Wikipedia, thus, provides a nice example of how epistemic values can come into conflict. In particular, while Wikipedia may be slightly less reliable than Encyclopedia Britannica, it is arguably much more powerful, speedy, and fecund. When there is such a conflict, we need to determine what the appropriate tradeoff is (cf. Fallis 2004a, 104, Fallis 2006, 500-503). And just as when reliability comes into conflict with non-epistemic interests, the relative importance of different epistemic values will often depend on the circumstances. For example, speed is often extremely important in our fast-paced world. It is sufficiently important to physicists that many of them use preprint archives that provide quick access to unpublished papers that have not been checked for accuracy by anyone other than the author (cf. Thagard 1997). Also, William James (1979 [1896], 31-32) famously claimed that the value of power can sometimes outweigh the value of reliability. According to James, "a rule of thinking which would absolutely prevent me from acknowledging certain kinds of truth if those kinds of truth were really there, would be an irrational rule." Therefore, in many circumstances, the epistemic benefits of Wikipedia (in terms

of greater power, speed, fecundity, and even verifiability) may very well outweigh the epistemic costs (in terms of somewhat less reliability).

Finally, as any user of Wikipedia knows (and as the empirical studies cited above suggest), it is not just a mass of misinformation and disinformation (cf. Sunstein 2006, 150). Wikipedia contains quite a lot of accurate high quality information. So, this is not simply a case of disseminating low quality information faster and to more people. Hence, despite legitimate concerns about its reliability, it probably is *epistemically* better (i.e., in terms of all of our epistemic values) that people have access to this information source.

8. Why Wikipedia is as Reliable as it is

According to Chris Anderson (2006, 71), "the true miracle of Wikipedia is that this open system of amateur user contributions and edits doesn't simply collapse into anarchy." As this quote suggests, an important task for social epistemologists is to explain why Wikipedia is as reliable as it is. In this section, I discuss the principal explanations that have been offered for the epistemic success of Wikipedia. None of these explanations is completely satisfying on its own, but collectively they do take us some way toward understanding the epistemic success of Wikipedia.

a) Wiki Technology

An obvious explanation is that any errors will be quickly found and corrected in Wikipedia because such a large number of people are working to remove them.[12] At the moment, approximately 75,000 people are active contributors to Wikipedia (see http://en.wikipedia.org/wiki/Wikipedia:About). And with wiki technology, any of these contributors can immediately correct any errors that they find in Wikipedia. The same sort of explanation is often given for the success of open source software projects (cf. Duguid 2006). In that instance, the slogan is "given enough eyeballs, all bugs are shallow."

Errors on the Internet can be corrected much more easily and quickly than errors in a printed book (cf. Thagard 1997). Instead of having to wait until the next edition is published, errors on the Internet can, at least in principle, be corrected *immediately*. Nevertheless, in actual practice, such corrections usually take somewhat longer because most websites can only be edited by a small number of people. By contrast, anyone who finds an error in Wikipedia can (and is encouraged to) immediately correct it herself. In addition, errors in Wikipedia are more likely to be found in the first place because more people are on the lookout for them. (Admittedly, many Wikipedia entries do not get a lot of traffic and, thus, will not be checked frequently for errors. But because they do not get a lot of readers, the potential epistemic cost of errors in these entries is correspondingly lower as well.)

Just as errors can be more easily corrected, however, they can also be more easily introduced (intentionally or unintentionally) into Wikipedia (cf. Duguid 2006). So, in order to explain why Wikipedia is as reliable as it is, we need an explanation for why errors are more likely to be corrected than introduced as a result of wiki technology.

b) The Wisdom of Crowds

A popular explanation of its reliability is that Wikipedia is an example of the *Wisdom of Crowds* (cf. Anderson 2006, 67-70, Sunstein 2006, 151-156). In his book of this title, James Surowiecki (2004) describes a number of examples of large groups that are extremely reliable. For instance, in 1906, visitors to a livestock exhibition in the west of England were asked to guess the weight of a fat ox in order to win a prize. The British scientist Francis Galton found that the average of the eight hundred guesses was within one pound of the actual weight (cf. Surowiecki 2004, xiii). Also, when contestants on *Who Wants to be a Millionaire?* need help with a question, they can call one of their smart friends for assistance or they can poll the studio audience. The studio audience gets the right answer approximately 91 percent of the time (cf. Surowiecki 2004, 4).

In fact, a large group will often be more reliable than any individual member of the group. For example, the smart friend only gets the right answer approximately 65 percent of the time. Also, the *Condorcet Jury Theorem* shows that, even if none of the voters in an election are all that reliable, they can still collectively be highly reliable (cf. Estlund 1994).

Nonetheless, not just any large group of people is going to be reliable on just any question. In order to be wise, a group must have certain properties. Surowiecki (2004, 10), for example, claims that wise groups must be large, independent, and diverse (cf. Page 2007, 175-235). Similarly, the Condorcet Jury Theorem is based on the assumption that there are a lot of voters, the voters are at least somewhat reliable, and their votes are independent.[13]

Wikipedia certainly has some of the properties that are associated with wise groups. For example, encyclopedias primarily collect factual information (such as the date that John Locke was born and the melting point of copper), which is the type of information that wise groups are especially likely to get right (cf. Surowiecki 2004). In fact, this may explain why Wikipedia seems to do somewhat better with scientific topics than with other topics which require more interpretation and analysis. Also, it might be suggested that people who voluntarily choose to write on a particular topic are likely to have at least some degree of reliability on that topic. Indeed, there are many amateur ornithologists, amateur astronomers, etc. who do not have academic credentials, but who have quite a bit of expertise. Nonetheless, as I discuss below, Wikipedia also lacks many of the properties that are associated with wise groups. Thus, it is not clear that appealing to the wisdom of crowds provides a sufficient explanation for why Wikipedia is as reliable as it is.

First, the examples of the wisdom of crowds typically involve a large number of individuals working on *one specific* problem. As noted above, a large number of people certainly do contribute to Wikipedia. But only a few of these contributors work on any given Wikipedia entry (cf. Lih 2004, Sunstein 2006, 152).

Second, the examples of the wisdom of crowds typically involve a large number of *diverse* individuals who bring different perspectives and knowledge to the problem. However, it is not clear how diverse the contributors to Wikipedia really are. As Roy Rosenzweig (2006, 127) points out, "*Wikipedia*'s authors do not come from a cross-section of the world's population. They are more likely to be English-speaking, males, and denizens of the Internet." And it is especially unclear how diverse the contributors to any specific entry are likely to be.

Finally, the examples of the wisdom of crowds typically involve simply *aggregating* the independent viewpoints of many individuals. For example, we might take an average of the individual viewpoints in the case of quantitative information (such as the weight of an ox) or we might take a majority vote in the case of qualitative information. But what a Wikipedia entry says is rarely determined by such aggregation mechanisms. Contributions to Wikipedia are added sequentially by single individuals. So, there is a sense in which the collective viewpoint is simply the viewpoint of whoever happened to be the last person to edit an entry before you looked at it (cf. Sunstein 2006, 158). And when there is any controversy on some particular issue, contributors to Wikipedia typically use the talk pages to try to reach a consensus rather than resort to a vote (see http://en.wikipedia.org/wiki/Wikipedia:Consensus). But this process is subject to small group dynamics that can make a group much less reliable (cf. Surowiecki 2004, 173-191, Page 2007, 213-214). In particular, an emphasis on reaching consensus can stifle dissent (i.e., eliminate the independence) that helps a group to be wise.

c) Wikipedia Policies

There are also several policies that contributors to Wikipedia are supposed to follow that clearly promote the reliability of Wikipedia. For example, as with traditional encyclopedias, all entries are supposed to be written from a *neutral point of view* (see http://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view). Among other things, this means that claims should only be included in an entry when there is consensus that they are true. When

there is no consensus, Wikipedia should not take a definitive position. Instead, contributors have to back off to the meta-level and only include claims about what different people *believe* about the issue.[14] This Neutrality policy is arguably conducive to reliability because consensus can often be "diagnostic of truth" (cf. Goldman 1999, 71).

Also, any claims that are likely to be challenged are supposed to be supported by references to *reputable published sources* (see http://en.wikipedia.org/wiki/Wikipedia:Verifiability). And contributors are encouraged to remove claims that are not supported by such sources. As a result, while Wikipedia does not have to bump up against reality directly, at least it has to bump up against the published literature, which provides some degree of quality control.

Of course, merely having such policies does not insure that contributors will actually obey them. But those contributors who have taken on greater administrative responsibilities do have some power to enforce these policies. For example, after Colbert made his suggestion about African Elephants, editing of this entry was blocked temporarily by such administrators to stop vandalism. Also, people are largely cooperative (cf. Surowiecki 2004, 119-126). They pretty much have to be in order for societies to function at all. For example, if people had to litigate all of the time to get contracts enforced, things would quickly grind to a halt. And, in fact, even in the absence of legal constraints, people regularly develop and abide by cooperative norms (cf. Ellickson 1991).

In addition, some Wikipedia policies provide incentives for contributors rather than placing constraints on them. For example, entries that meet certain standards of quality may be selected by the contributors to Wikipedia to be *Featured Articles* (see http://en.wikipedia.org/wiki/Wikipedia:Featured_articles). And the prospect of having one's work advertised and recognized in this way can certainly inspire contributors to try to meet these standards of quality (which includes accuracy). Even contributors who remain anonymous can try to gain a reputation within the Wikipedia community under their user names (cf. endnote 12).

In general, contributors who strive for quality and accuracy are more likely to see their changes persist, to have their articles featured, and to gain such a reputation.

9. How Wikipedia can be Improved

Whatever the complete explanation for it might be, the empirical evidence does indicate that Wikipedia is fairly reliable. Moreover, even if there were evidence that Wikipedia was seriously unreliable, it is unlikely that we could convince the owners to shut it down or convince everybody else not to use it. Given that Wikipedia seems to be here to stay, what social epistemologists and information scientists can try to do is figure out how to improve the situation.

Library and information scientists have expended a lot of effort teaching people how to evaluate the quality of information that they find on the Internet (cf. Fallis 2004b, 477-478). However, in addition to trying to improve how people use Wikipedia, we can try to improve Wikipedia itself. For instance, as noted above, there are new technologies (e.g., Wikiscanner) that have the potential to increase the reliability and verifiability of Wikipedia. Also, there are new policies (e.g., contributors are no longer allowed to claim credentials unless they can prove that they actually have them) that are designed to do the same (cf. Read 2007b). Such developments fall under the so-called *ameliorative* project in epistemology which focuses on how we can modify our institutions and practices to better achieve our epistemic goals (cf. Kitcher 1992, 64).

In addition to improving Wikipedia, social epistemologists can also try to figure out how to improve *on* Wikipedia. For example, Larry Sanger is working to create a more reliable alternative to Wikipedia (see http://en.citizendium.org/wiki/CZ:About). *Citizendium.org* welcomes experts to contribute as authors and as editors, and entries that meet certain standards of quality are officially "approved" by such qualified editors. (Also, contributors to Citizendium are not allowed to remain anonymous.) In addition, *Veropedia.com* is an attempt to create a more reliable extension of Wikipedia. This website will host stable versions of Wikipedia entries that

have been approved by experts in the relevant subject areas (see http://veropedia.com/docs/faq.php). These projects are in line with Ross Atkinson's (1996, 254-257) suggestion that, while libraries and other information services should provide access to as many materials as possible, they should make the materials that are easiest to access as reliable as possible.

Even so, it is important to keep in mind that any proposed changes to Wikipedia are likely to have epistemic costs as well as epistemic benefits. For example, if we try to improve its reliability by giving experts more editorial control, we might end up decreasing its power since other people might be deterred from contributing (cf. Giles 2007). Also, if we try to improve its reliability by only including entries approved by experts, we might end up decreasing its speed since it will take longer to add and update entries. So, in order to evaluate any proposed changes, we need to be clear about exactly what our epistemic values are and what the appropriate tradeoffs are when there are conflicts.

In closing, it is worth noting that, while many people who want to improve Wikipedia focus on increasing its reliability, this is a fairly difficult task. For example, as noted above, it is not easy to get experts involved and to do so in a way that does not have other epistemic costs. Also, there is a lot of inertia to overcome if the initial draft of an entry is of very low quality. As a result, we might do better at improving the epistemic consequences of Wikipedia by *increasing its verifiability* (cf. Fallis 2004b, 477-478).

Ultimately, it is the readers who have to decide whether to believe what they read on Wikipedia. And, because Wikipedia entries are of such uneven quality, it is especially important that readers be able to verify the accuracy of this information. But there are all sorts of things that can feasibly be done to make this job easier for them (cf. Fallis 2004b, 478-481).

In order to increase the verifiability of Wikipedia, we need to provide readers with easy access to evidence of the quality of particular entries.[15] For example, while it can take a lot of effort for experts to create new content for Wikipedia or even just correct existing content, it is

fairly easy to simply flag existing content that is questionable. In fact, it might be useful to set up a mechanism that allows people in general to rate the quality of the content as they do on *Digg.com* or *Amazon.com*. Such a mechanism would get readers, even those who do not actually contribute content to a specific entry, involved in making the entry more verifiable. And, *pace* Keen (2007, 6), this is the sort of application where the wisdom of crowds does seem to work.

Furthermore, it would presumably be fairly easy for experts to direct readers to good introductory sources on any given topic. For example, since Wikipedia is where many people start their research, libraries and archives with good digital materials on various topics are putting links to these materials in the relevant Wikipedia entries (cf. Lally and Dunford 2007). Also, Wikipedia is clearly more verifiable if it has more references to authoritative sources. And it is even easier for people to verify the information on Wikipedia if those authoritative sources are also freely available on the Internet. Thus, open access peer-reviewed journals can provide a useful supplement to Wikipedia (cf. Willinsky 2007).

It would also be useful to know which *types* of entries are more likely to be questionable. Such knowledge would allow us to be more cautious when consulting Wikipedia on certain matters. Also, it would tell us when we might be better served consulting other information sources in the first place. But so far, most empirical studies of the reliability of Wikipedia have looked at entries in one specific subject area (e.g., science, medicine, history, philosophy). It would be useful to carry out some empirical studies that compare the reliability of Wikipedia with respect to different subject areas. For example, it would be good to confirm the plausible hypothesis that Wikipedia does better with scientific topics and current events than it does with philosophical topics.

Finally, several empirical studies (e.g., Bragues 2007, Devgan et al. 2007) suggest that the more serious problem with Wikipedia is not that entries are likely to be inaccurate, but that they are likely to be incomplete.[16] People can be misled by incomplete information as well as inaccurate information (cf. Frické and Fallis 2004, 240). Even if Wikipedia includes only

accurate information, if its omissions tend to mislead people, it may not be a *reliable* information source in the sense that it does not lead to a high percentage of true beliefs relative to the total number of beliefs formed (cf. Thagard 1997). (This is another reason why people should be reminded that encyclopedias in general should be a starting point rather than an ending point for research.) Hence, important omissions should be flagged as well as inaccuracies. In addition, those pundits who want to warn people about the dangers of Wikipedia should probably not be focusing on the (alleged) inaccuracy of the information (e.g., by equating Wikipedia with public restrooms or infinite numbers of monkeys), but rather on the incompleteness of the information.

10. Conclusion

Like the Internet itself, Wikipedia is having a huge impact on how a great many people gather information about the world. So, it is important for epistemologists to ask what the epistemic consequences are of people having access to this information source. While there are legitimate concerns about its reliability (since anyone can edit it), the empirical evidence suggests that Wikipedia is fairly reliable (especially compared to those information sources that are as easily accessible). In addition, it has a number of other epistemic virtues (e.g., power, speed, and fecundity) that arguably outweigh any deficiency in terms of reliability. Even so, epistemologists should be trying to identify changes (or alternatives) to Wikipedia that will bring about even better epistemic consequences. And, in order to do that, we need to clarify what our epistemic values are and we need a better understanding of why Wikipedia works as well as it does.

11. Acknowledgements

An earlier version of this paper was part of a panel presentation on Wikipedia at the *North American Computing and Philosophy Conference*. I would like to thank the audience and the other panelists (Tony Doyle and Marc Meola) for helpful comments and suggestions. I would like to thank Julia Annas, Martin Frické, Bruce Fulton, Alvin Goldman, Rachana Kamtekar, Peter Lewis, P. D. Magnus,

Kay Mathiesen (and the students in her graduate seminar on *Distributed Cognition* at the University of Arizona), Heshan Sun, Dennis Whitcomb, K. Brad Wray, and the students in my course on *Social Epistemology and Information Science* at the University of Arizona for their feedback. I would like to thank the anonymous reviewers for JASIST for several suggestions that have significantly improved this paper. (A slightly different version of the first half of this paper is expected to appear, along with a piece by Larry Sanger, as "Wikipistemology: A Debate" in Alvin Goldman and Dennis Whitcomb (Eds.), *Social Epistemology: An Anthology*, Oxford University Press, 2008.)

12. References

Adler, B. T., & de Alfaro, L. (2007). A content-driven reputation system for the Wikipedia. *Proceedings of the International World Wide Web Conference*. Retrieved November 15, 2007 from http://www.soe.ucsc.edu/~luca/papers/07/wikiwww2007.pdf

Anderson, C. (2006). *The long tail*. New York: Hyperion.

Anthony, D., Smith, S. W., & Williamson, T. (2007). The quality of open source production: Zealots and good Samaritans in the case of Wikipedia. *Dartmouth Technical Report*. Retrieved November 15, 2007 from http://www.cs.dartmouth.edu/reports/abstracts/TR2007-606/

Atkinson, R. (1996). Library functions, scholarly communication, and the foundation of the digital library: Laying claim to the control zone. *Library Quarterly, 66*, 239-265.

Black, M. (1983). *The prevalence of humbug and other essays*. Ithaca: Cornell University Press.

Birnholtz, J. P. (2007). When do researchers collaborate? Toward a model of collaboration propensity. *Journal of the American Society for Information Science and Technology, 58*, 2226-2239.

Borland, J. (2007). See who's editing Wikipedia: Diebold, the CIA, a campaign. *Wired*. Retrieved November 15, 2007 from http://www.wired.com/politics/onlinerights/news/2007/08/wiki_tracker

Bragues, G. (2007). Wiki-philosophizing in a marketplace of ideas: Evaluating Wikipedia's entries on seven great minds. Retrieved November 15, 2007 from http://ssrn.com/abstract=978177

Coady, D. (2006). When experts disagree. *Episteme, 3*, 68-79.

Cohen, G. A. (2002). Deeper into bullshit. In S. Buss, & L. Overton (eds.), *Contours of agency* (pp. 321-339). Cambridge: MIT University Press.

Cohen, N. (2007). Courts turn to Wikipedia, but selectively. *New York Times*. Retrieved November 15, 2007 from http://www.nytimes.com/2007/01/29/technology/29wikipedia.html

Cronin, B. (2005). BLOG: see also Bathetically Ludicrous Online Gibberish. Retrieved November 15, 2007 from http://www.slis.indiana.edu/news/story.php?story_id=958

Cross, T. (2006). Puppy smoothies: Improving the reliability of open, collaborative wikis. *First Monday, 11*. Retrieved February 1, 2008 from http://www.firstmonday.org/ISSUES/issue11_9/cross/

Denning, P., Horning, J., Parnas, D., & Weinstein, L. (2005). Wikipedia risks. *Communications of the ACM, 48*, 152.

Descartes, R. (1996). *Meditations on first philosophy*. Cambridge: Cambridge University Press.

Devgan, L., Powe, N., Blakey, B., & Makary, M. (2007). Wiki-Surgery? Internal validity of Wikipedia as a medical and surgical reference. *Journal of the American College of Surgeons, 205*, S76-S77.

Duguid, P. (2006). Limits of self-organization: Peer production and "laws of quality". *First Monday, 11*. Retrieved November 15, 2007 from http://www.firstmonday.org/issues/issue11_10/duguid/

Ellickson, R. C. (1991). *Order without law*. Cambridge: Harvard University Press.

Encyclopaedia Britannica. (2006). Fatally flawed: Refuting the recent study on encyclopedic accuracy by the journal *Nature*. Retrieved November 15, 2007 from http://corporate.britannica.com/britannica_nature_response.pdf

Estlund, D. M. (1994). Opinion leaders, independence, and Condorcet's jury theorem. *Theory and Decision, 36*, 131-162.

Fallis, D. (2004a). Epistemic value theory and information ethics. *Minds and Machines, 14*, 101-117.

Fallis, D. (2004b). On verifying the accuracy of information: Philosophical perspectives. *Library Trends, 52*, 463-487.

Fallis, D. (2006). Social epistemology and information science. In B. Cronin (Ed.), *Annual review of information science and technology* (Vol. 40, pp. 475-519). Medford, New Jersey: Information Today.

Fallis, D. (2007). Toward an epistemology of intellectual property. *Journal of Information Ethics, 16*, 34-51.

Fallis, D., & Frické, M. (2002). Indicators of accuracy of consumer health information on the Internet. *Journal of the American Medical Informatics Association, 9*, 73-79.

Feldman, R. (2003). *Epistemology*. Upper Saddle River, New Jersey: Prentice Hall.

Fox, C., Levitin, A., & Redman, T. (1994). The notion of data and its quality dimensions. *Information Processing and Management, 30*, 9-19.

Frankfurt, H. G. (2005). *On bullshit*. Princeton: Princeton University Press.

Frické, M., & Fallis, D. (2004). Indicators of accuracy for answers to ready reference questions on the Internet. *Journal of the American Society for Information Science and Technology, 55*, 238-245.

Giles, J. (2005). Internet encyclopaedias go head to head. *Nature, 438*, 900-901.

Giles, J. (2007). You can put your trust in Wikipedia 2.0: Wikipedia wants to be seen as more reliable, but will the changes it is planning deter genuine contributors? *New Scientist, 195*, 28-29.

Goldman, A. I. (1999). *Knowledge in a social world*. New York: Oxford University Press.

Goldman, A. I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research, 63*, 85-110.

Goldman, A. I. (2008). The social epistemology of blogging. In J. van den Hoven, & J. Weckert (eds.), *Information technology and moral philosophy*. Cambridge: Cambridge University Press. Retrieved January 17, 2008 from http://fas-philosophy.rutgers.edu/goldman/Social%20Epistemology%20of%20Blogging.pdf

Gorman, G. E. (2007). A tale of information ethics and encyclopedias; or, is Wikipedia just another internet scam? *Online Information Review, 31*, 273-276.

Guinnee, E. (2007). A new context for knowledge creation. *Library Student Journal*. Retrieved

November 15, 2007 from http://www.librarystudentjournal.org/index.php/lsj/article/view/18/44

Hardin, R. (2003). If it rained knowledge. *Philosophy of the Social Sciences, 33*, 3-24.

Hernon, P. (1995). Disinformation and misinformation through the internet: Findings of an exploratory study. *Government Information Quarterly, 12*, 133-139.

Hume, D. (1977). *An enquiry concerning human understanding*. Indianapolis: Hackett.

Impicciatore, P., Pandolfini, C., Casella, N., & Bonati, M. (1997). Reliability of health information for the public on the world wide web: systematic survey of advice on managing fever in children at home. *British Medical Journal, 314*, 1875-1879.

James, W. (1979). *The will to believe and other essays in popular philosophy*. Cambridge: Harvard University Press.

Keen, A. (2007). *Cult of the amateur*. New York: Doubleday.

Kelton, K., Fleischmann, K. R., & Wallace, W. A. (2008). Trust in digital information. *Journal of the American Society for Information Science and Technology, 59*, 363-374.

Kitcher, P. (1992). The naturalists return. *Philosophical Review, 101*, 53-114.

Kronick, D. A. (1988). Anonymity and identity: editorial policy in the early scientific journal. *Library Quarterly, 58*, 221-237.

Lackey, J., & Sosa, E. (2006). *The epistemology of testimony*. Oxford: Oxford University Press.

Lally, A. M., & Dunford, C. E. (2007). Using Wikipedia to extend digital collections. *D-Lib Magazine, 13*. Retrieved November 15, 2007 from http://www.dlib.org/dlib/may07/lally/05lally.html

Levi, I. (1967). *Gambling with truth*. Cambridge: MIT Press.

Lih, A. (2004). Wikipedia as participatory journalism: Reliable sources? Metrics for evaluating collaborative media as a news resource. *Proceedings of the International Symposium on Online Journalism*. Retrieved November 15, 2007 from http://jmsc.hku.hk/faculty/alih/publications/utaustin-2004-wikipedia-rc2.pdf

Luyt, B., Aaron, T. C. H., Thian, L. H., & Hong, C. K. (2008). Improving Wikipedia's accuracy: Is edit age a solution? *Journal of the American Society for Information Science and Technology, 59*, 318-330.

Magnus, P. D. (2006). Epistemology and the Wikipedia. Paper presented at the *North American Computing and Philosophy Conference*. Retrieved November 22, 2007 from http://hdl.handle.net/1951/42589

Maher, P. (1993). *Betting on theories*. New York: Cambridge University Press.

Mann, T. (1993). *Library research models*. New York: Oxford University Press.

Meola, M. (2000). Review of *Knowledge in a social world* by Alvin I. Goldman. *College and Research Libraries, 61*, 173-174.

Meyer, B. (2006). Defense and illustration of Wikipedia. Retrieved November 15, 2007 from http://se.inf.ethz.ch/~meyer/publications/wikipedia/wikipedia.pdf

Mill, J. S. (1978). *On liberty*. Indianapolis: Hackett.

Page, S. E. (2007). *The difference*. Princeton: Princeton University Press.

Pang, A. S.-K. (1998). The work of the encyclopedia in the age of electronic reproduction. *First Monday, 3*. Retrieved November 15, 2007 from http://www.firstmonday.org/issues/issue3_9/pang/

Paterson, R. W. K. (1979). Towards an axiology of knowledge. *Journal of Philosophy of Education, 13*, 91-100.

Plato. (1961). *Collected Dialogues*. Princeton: Princeton University Press.

Poe, M. (2006). The Hive. *Atlantic Monthly*. Retrieved November 15, 2007 from http://www.theatlantic.com/doc/200609/wikipedia

Pritchard, D. (2007). Recent work on epistemic value. *American Philosophical Quarterly, 44*, 85-110.

Rainie, L., & Tancer, B. (2007). Data memo. Retrieved November 15, 2007 from http://www.pewinternet.org/pdfs/PIP_Wikipedia07.pdf

Read, B. (2006). Can Wikipedia ever make the grade? *Chronicle of Higher Education, 53*(10), A31.

Read, B. (2007a). Middlebury College History Department limits students' use of Wikipedia.

Chronicle of Higher Education, 53(24), A39.

Read, B. (2007b). Wikipedia fights bogus credentials. *Chronicle of Higher Education*. Retrieved November 15, 2007 from http://chronicle.com/wiredcampus/article/1923/expertise-on-trial-on-wikipedia

Rieh, S. Y. (2002). Judgment of information quality and cognitive authority in the Web. *Journal of the American Society for Information Science and Technology, 53*, 145-161.

Riggs, W. D. (2003). Balancing our epistemic ends. *Nous, 37*, 342-352.

Rosenzweig, R. (2006). Can history be open source? *Wikipedia* and the future of the past. *Journal of American History, 93*, 117-146.

Sanger, L. (2005). The early history of Nupedia and Wikipedia: A memoir. Retrieved November 15, 2007 from http://features.slashdot.org/article.pl?sid=05/04/18/164213

Schiff, S. (2006). Know it all. *New Yorker*. Retrieved November 15, 2007 from http://www.newyorker.com/archive/2006/07/31/060731fa_fact

Shera, J. (1970). *Sociological foundations of librarianship*. New York: Asia Publishing House.

Sunstein, C. R. (2006). *Infotopia*. Oxford: Oxford University Press.

Surowiecki, J. (2004). *The wisdom of crowds*. New York: Doubleday.

Swanson, D. R. (1986). Fish Oil, Raynaud's syndrome, and undiscovered public knowledge. *Perspectives in Biology and Medicine, 30*, 7-18.

Tapscott, D., & Williams, A. D. (2006). *Wikinomics*. New York: Portfolio.

Thagard, P. (1997). Internet epistemology: Contributions of new information technologies to scientific research. Retrieved November 15, 2007 from http://cogsci.uwaterloo.ca/Articles/Pages/Epistemology.html

Thagard, P. (2006). How to collaborate: Procedural knowledge in the cooperative development of science. *Southern Journal of Philosophy, 44*, S177-S196.

Tollefsen, D. (2007). Group testimony. *Social Epistemology, 21*, 299-311.

Viegas, F. B., Wattenberg, M., & Dave, K. (2004). Studying cooperation and conflict between authors with *history flow* visualizations. *Conference on Human Factors in Computing Systems*.

Retrieved November 15, 2007 from http://alumni.media.mit.edu/~fviegas/papers/history_flow.pdf

Waldron, J. (1995). The wisdom of the multitude: Some reflections on Book 3, Chapter 11 of Aristotle's *Politics*. *Political Theory, 23*, 563-584.

Warner, J. (2005). Labor in information systems. In B. Cronin (Ed.), *Annual review of information science and technology* (Vol. 39, pp. 551-573). Medford, New Jersey: Information Today.

Willinsky, J. (2007). What open access research can do for *Wikipedia*. *First Monday, 12*. Retrieved January 17, 2008 from http://www.firstmonday.org/issues/issue12_3/willinsky/

Wilson, P. (1983). *Second-hand knowledge*. Westport, Connecticut: Greenwood Press.

Winchester, S. (1998). *The professor and the madman*. New York: HarperCollins.

Wray, K. B. (2002). The epistemic significance of collaborative research. *Philosophy of Science, 69*, 150-168.

13. Endnotes

---

[1] There is a sense in which science, history, and even language itself are the result of mass collaboration. But it is now possible for many people to collaborate on a single, narrowly-defined project over a relatively short span of time.

[2] Information scientists sometimes claim that knowledge "has nothing to do really with truth or falsehood" (Shera 1970, 97). But there is a significant amount of legitimate concern with accuracy and reliability in information science (cf. Meola 2000). In this paper, I follow most contemporary epistemologists and assume that knowledge is some form of justified true belief (cf. Goldman 1999, Feldman 2003). See Fallis (2006, 490-495) for a defense of using this definition of knowledge within information science.

[3] In addition to creating and editing content, contributors to Wikipedia can (with the approval of a sufficient number of their peers) also take on greater administrative responsibilities. For example, such contributors might be charged with monitoring entries for vandalism, moderating disputes between contributors, and even restricting the editing of certain entries when necessary. See http://en.wikipedia.org/wiki/Wikipedia:About for further details about how Wikipedia works.

[4] Of course, as a direct result of what they have learned from encyclopedias, people may go on to discover new knowledge. In general, collecting a lot of existing knowledge in one place is often a good method for making new discoveries (cf. Swanson 1986).

[5] Another concern that has been raised is that the content of Wikipedia is constantly in flux. For example, how can you cite a Wikipedia entry to support a point that you want to make if that entry might say something totally different tomorrow? However, constant flux is really the price that we always have to pay if an information source is to be up-to-date. Also, since every version of every entry is archived and accessible, it is possible to cite the specific version of the Wikipedia entry that supports the point that you want to make.

[6] In addition to people who intentionally add inaccurate or misleading information, there are people who intentionally remove content and/or replace it with obscenities (see http://en.wikipedia.org/wiki/Wikipedia:Vandalism). If it removes *true* content, such vandalism also reduces the reliability of Wikipedia.

[7] Of course, some bullshit might turn out to be true. In fact, some information that is intended to be false might accidentally turn out to be true. And, if it is true, then there is not much of an epistemic cost to having it in Wikipedia. However, it seems safe to say that any given instance of bullshit or disinformation is unlikely to true. For example, it seems rather unlikely that the number of African Elephants has recently tripled.

[8] Encyclopedia Britannica (2006) has criticized the methodology of this study. But *Nature* has defended its methodology (see http://www.nature.com/nature/britannica/). The bottom line is that there is no reason to think that any methodological failings of the study would favor Wikipedia over Britannica (cf. Magnus 2006). See http://en.wikipedia.org/wiki/Reliability_of_Wikipedia for an extensive survey of empirical studies and expert opinions on the reliability of Wikipedia.

[9] Strictly speaking, the websites returned by a search engine are not randomly chosen. In fact, Google does tend to return more accurate websites higher in its search results than less accurate websites (cf. Frické and Fallis 2004, 243).

[10] Such word-for-word copies are probably the most likely non-discriminating reflectors of Wikipedia. But as P. D. Magnus has pointed out to me, a website might use Wikipedia as a source without copying

Wikipedia word-for-word and without citing Wikipedia. In such cases, it will be difficult for readers to determine that they are dealing with a non-discriminating reflector. But it is interesting to note that even non-discriminating reflectors can provide some corroboration for the information on Wikipedia. As David Coady (2006) points out, if we have reason to believe that someone is good at identifying reliable sources, then the fact that she has chosen to be a non-discriminating reflector of a particular source is some evidence that that source is reliable.

[11] Cross (2006) has claimed that *how long* a piece of text has survived in a Wikipedia entry is a good indication that this text is accurate. Text could be color-coded to make this indicator of accuracy easily accessible to readers. However, subsequent research (viz., Lutz et al. 2008) suggests that edit age is not an indicator of accuracy.

[12] Since the number of contributors seems to have something to do with the reliability of Wikipedia, we really also need to ask why so many people voluntarily choose to contribute for free. Anthony et al. (2007) suggest that many people are interested in gaining a reputation within the Wikipedia community. For example, since all of your edits can be tagged to your user name, you can strive to get on the list of Wikipedia contributors with the most edits (see http://en.wikipedia.org/wiki/Wikipedia:List_of_Wikipedians_by_number_of_edits). But for purposes of this particular paper, I am just going take it for granted that many people *are* motivated to contribute.

[13] It turns out that it is diversity rather than independence that makes a group wise (cf. Estlund 1994, Page 2007, 202). In fact, a group is wiser if the views of its members are negatively correlated. (Independence is just a way to insure that there is a reasonable degree of diversity.)

[14] Another concern about Wikipedia is that contributors who disagree about a particular issue will keep changing an entry back and forth. Such "edit wars" make Wikipedia less stable as well as less reliable. But adherence to the Neutrality policy can help to avoid such disputes (cf. Sanger 2005).

[15] Methods of verifying claims in different subject matters can differ. For example, the appropriate empirical tests for claims in physics will be different from the tests that are appropriate for claims in psychology. But readers will rarely be able to directly check the claims made on Wikipedia. For example, few people are in a position to check the melting point of copper for themselves. It is usually much more

feasible to check the claims indirectly (e.g., by checking to see if the source might be biased or if other sources corroborate the claim). Such indirect techniques tend to be independent of the subject matter.

[16] Wikipedia provides much more information, and often much more detail, than traditional encyclopedias. But relative to this greater level of detail, there are apparently significant gaps. Also, it should be noted that such incompleteness may not be the only way in which accurate information on Wikipedia can potentially leave a false impression.