

3. Should you believe Wikipedia?

In the previous chapter, I explored how Wikipedia works. But is it any good? Should you believe what you read on Wikipedia? My children’s middle and high-school teachers all forbid students from citing it, and some ask that they not use it at all. Are they right? To answer that question, we need first to back up and ask a bigger question: how do we know anything at all?

How Do You Know?

Do you believe that human activity is changing our global climate? One way or the other, how do you know? I walked around my department one morning asking that question. One person said she knew climate change was real because “scientific data supports it.” I asked, “Have you actually read any papers on the topic?” Paraphrasing, here are the answers I got:

- “I’ve read abstracts of a few scientific studies and heard reports on the science on news sources I trust.”
- “I read one study all the way through, and beyond that heard about it on the news.”
- “I’ve never actually read any climate science, but I’ve heard news reports.”
- “They taught us about it in sixth grade.”
- “I saw Al Gore’s movie.”
- “I took a trip to Alaska recently. We had to walk two hours to get to the glacier. They told me that just a few years ago, the glacier was right there—you didn’t have to walk to get to it at all.”

All my colleagues believe climate change is real. None of them are actually climate scientists. In the absence of enough time to really study the issue, they rely on sources they trust. But what about the other side—the people who do not believe that human activity is changing our climate? How have they come to that conclusion?

For the sake of argument, let’s take as a given that human activity is really changing the climate—climate change is real. Is the other side simply delusional? Have we entered the “post truth” era? “Post truth” was the word of the year for 2016, according to the Oxford Dictionaries. It is defined as “relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief” (Word of the Year 2016). This perspective implies there are thoughtful people who rely on facts, and careless people who don’t. It’s certainly true that some people are intellectually careless (we can’t tell how many), but I will argue that this view fails to describe what is taking place in a useful way. It’s not that one side is relying on sources and the other is not—both sides have sources. Unless you are a true expert on a given topic, the best you can do is rely on sources you trust. The problem then becomes: *How do you decide which sources to trust?*

Information overload and the complexity of modern life mean that we don’t have time to make truly informed decisions on most issues. I don’t have time to get a graduate degree in either climate science or immunology—I need quicker ways than that to decide whether it is worth paying extra money to buy a low emissions car, and whether it is safe to vaccinate my children. So I rely on sources I trust.

Each person's decisions about which sources to trust are somewhat mysterious. In a bygone era, people were told what to believe on important matters by community and religious leaders. In contemporary times those decisions are often shaped by mass media (starting with the inventions of print, radio, and television) and more recently by the internet and social media.

If we now assume that climate change is debatable, how can we have that debate? For example, suppose we ask, "Should I believe Al Gore's movie?" (An Inconvenient Truth n.d.). The next logical step would be to find the sources that support the film. But then we need to find the sources for those sources—and on ad infinitum. Is there any kind of "base fact" that is beyond dispute, that we can then build on? This is a question of *epistemology*, the theory of knowledge. If we now ask, "Why do some people choose to believe Al Gore's movie, and others don't?" that is a question of *sociology of knowledge*.

On most issues, people are predisposed to believe a certain way, and therefore accept sources that agree. This is called "confirmation bias" (Nickerson 1998). On the internet, there are sources to support every possible point of view. If you believe that humanity is arrogant and we have not been good stewards of our natural resources, then you are more inclined to believe that climate change is real and read The Huffington Post. If you believe that humanity has been divinely given dominion over the earth and natural resources are there expressly for our use, then you are more inclined to believe that climate change is not real and read Breitbart. Each world view has a set of ideas that are easily assimilated and ideas that don't seem to fit. The internet helps create new world views, with sources to support them and other people who believe them. Those groups—whether they have typical or atypical views—support one another's beliefs.

Those worlds are increasingly isolated. When I walked around my office building asking about climate change, I couldn't find a single person who had doubts. Or if they did, they were not comfortable admitting it. Social worlds foster views of reality, and those social worlds these days sometimes don't overlap in any harmonious way.

People who do and do not believe in climate change use similar methods to decide what to believe—media and human sources they trust. But that does not mean that both sides are "right." A preponderance of evidence suggests climate change is real—there is a "true" answer. How do I know? Because I am willing to assert that my sources (high-quality journalism and peer-reviewed scientific publications) are more robust than competing sources.

What does it mean for something to be "true"? How is the internet changing how we understand truth? This chapter explores how theories of the nature of truth and knowledge can help us to understand the internet. Of course, one could spend a lifetime trying to truly understand these questions (and many people do). Here I will simply give a high-level introduction to parts of these theories I have found helpful as an internet user and researcher.

Metaphysics: Internal Realism

Everything we know, we experience through our senses. Our senses are fallible. We can make mistakes in what we perceive, and every observation has a limit to its accuracy.

With our fallible senses, we create representations of the world—both internal (mental) ones and external ones (writing, maps, etc.). Those representations are always approximate. When I was studying high-school physics, I thought that we were working with approximate models of situations, because we were just in high school—later we would learn the “real” solutions. As an undergraduate physics major, I quickly learned that all solutions would always be approximations. The models get more and more sophisticated, but they are always approximations. As Jorge Luis Borges wrote, the only perfect map of the territory is the territory itself (Borges and Hurley 1999). The universe is, atom for atom, a perfect model for itself—and anything else has some degree of error in representation.

To further complicate things, the act of observation changes the thing being measured. Sometimes that difference is significant, and sometimes not. It’s clearly significant in the subatomic realm, where bouncing a photon off of a particle to observe it moves the particle. It’s also clear in the social realm, where for example the presence of an observer shapes an individual’s behavior. Erving Goffman wrote about how humans are always performing for others, trying to make a particular impression (Goffman 1959). The impressions given (what we intend to communicate) may be different from the impressions given off (what others actually infer from our performance) (more on this in Chapter 5). In many other cases, the act of observation is less significant.

If all I know is what my senses tell me, how do I know I’m not dreaming? How do I know that the world exists at all? This is a question of metaphysics. An objectivist view argues that the world exists and we can reliably know things about it. A subjectivist view argues that we are trapped within the limitations of our senses and can not definitively know anything.

Much ink has been spilled debating objectivist versus subjectivist views of reality. Fortunately, there is a common-sense compromise called “internal realism.” How do I know the world exists? Honestly, I don’t. But statistically, it’s highly likely. I am trapped within my subjective perceptions and the limitations of my senses. You are trapped within yours. But let’s suppose that we both agree that I am sitting on a chair. If our perceptions are so fallible, why do we agree? Now let’s suppose I ask the hundred or so people who work in this building to come by and offer their opinions, and everyone agrees that I am sitting on a chair. The reason that we all agree is that *the chair exists*. I can’t prove it, but the high degree of correlation between our subjective perceptions leads me to believe that the world is real, independent of my perceptions of it. This philosophical approach was first articulated by Hilary Putnam (Lakoff 1987). The world exists, but is only knowable through the fallible senses of people who are part of the world (not separate from it). The limits of our senses and the fact that we are part of the world we are observing both lead to limits on the accuracy of what we can claim “is true.”

Epistemology

The world exists, but through our limited point of view and fallible senses, what can we know about it? What does it mean to know something? Epistemology is the study of knowledge. To know something,

you must believe it, it must be true, and your belief must be justified (Steup 2016). If I am a contestant on a game show and say, “I know the grand prize is behind door number three!”, that is not knowledge, even if I happen to be right. My belief (though correct) was not justified. On the other hand, if I am a stage hand and saw the grand prize behind door number three, then I have knowledge (true, justified belief) that the prize is there. I know because I saw it with my senses. One candidate for a justified belief is something we directly perceive with reliable senses.

How do we decide whether a belief is justified? Epistemologists debate the nature of justification, and whether justification is necessary or makes sense at all. Two basic approaches to justification are evidentialism and reliabilism. Evidentialism says beliefs must be supported by evidence--all beliefs can be derived from a set of justified, basic beliefs. Basic beliefs include things you learn through reliable senses, and spontaneously formed beliefs if they are a “proper response to experience” (Feldman 2003). Nonbasic beliefs are formed by inference from basic ones. In this view, it’s challenging to know anything if we have to follow a chain of evidence all the way down to basic beliefs.

On the other hand, reliabilism suggests that a belief is reliable if it is the result of a reliable cognitive process. A reliable process is one that generally leads to correct outcomes. Richard Feldman provides the example of two bird watchers:

“Compare a novice bird-watcher and an expert walking together in the woods, seeking out the rare pink-spotted flycatcher. A bird flies by and each person spontaneously forms the belief that there is a pink-spotted flycatcher there. The expert knows this to be true but the novice is jumping to a conclusion out of excitement. The expert has a well-founded belief but the novice does not.” (Feldman 2003)

Having a truly reliable process can be a challenge. The theory of *virtue epistemology* argues that obtaining knowledge is *an achievement*. Knowledge is the result of “competent cognitive agency” (Greco 2021). Epistemic virtues include “curiosity, intellectual autonomy, intellectual humility, attentiveness, intellectual carefulness, intellectual thoroughness, open-mindedness, intellectual courage and intellectual tenacity” (Heersmink 2017). Critically, obtaining knowledge is not usually something you do by yourself—it is a social process. Knowledge is the result of successful collaboration.

When are you justified in believing something because someone else told you? In epistemology, this is the problem of *testimony*. By accepting someone’s testimony, you make yourself vulnerable to possibly accepting bad information either because of their ignorance or dishonesty. The problem of testimony comes down to deciding who to trust.

If testimony that you accept is correct, you are collaboratively building knowledge. To explain how testimony is collaborative, John Greco draws an analogy to a soccer pass:

“Soccer: Playing in a soccer game, Ted receives a brilliant, almost impossible pass. With the defense out of position and goalie lying prostrate on the ground, Ted kicks the ball into the net for an easy goal.” (Greco 2021)

The goal is a collaborative achievement between the midfielder and our striker, Ted. In this example, most of the credit goes to the midfielder, but it’s definitely a collaborative effort. Greco argues that gaining knowledge through testimony is similar. If all goes well, it is a collaborative achievement of the speaker and the listener. Further, the social environment shapes whether successful knowledge transmission is likely to take place. Greco writes that, “In cases of knowledge transmission, the reliability of the testimonial exchange is at least partly explained by the social environment. More exactly, it is explained (in part) by the social norms that structure the social environment, and the ways hearer and speaker sensibilities interact with those norms” (Greco 2021).

These ideas from epistemology give us new ways of understanding the internet. When we receive information online, we can ask: am I justified in accepting this as true? Justified because of evidence I have direct access to or because it comes from a reliable source?

Receiving information from someone else online (like a post on social media) is an attempt at creating knowledge through testimony. Reliability of the testimonial exchange is determined by features of the environment (the social media platform), and social norms for people contributing content there. The design of the platform shapes whether the result is likely to be knowledge (true, justified belief) or not. Internet users can pay attention to features of platforms to help them decide what to believe. Designers of platforms have an important job: to design to facilitate the successful creation of knowledge.

Social Construction of Knowledge

The fact that everyone in the Technology Square Research Building agrees that I am sitting on a chair led me to the qualified statement that it is highly likely that there is such a thing as “a chair” and I am sitting on one. More broadly, that’s a pretty good metaphor for how science works. In 1908, Robert Millikan and Harvey Fletcher did an experiment with oil drops and concluded that the charge on an electron is unitary—depending on how many electrons you have, the charge is all in multiples of the same base number (Holton 1978). At the time, that was a controversial claim. Today, we would say “Millikan was right.” How do we know? We know because more than a hundred years of subsequent scientific research confirms that finding.

Millikan made some observations and drew conclusions—but he might’ve been wrong. In fact, his experiment was controversial because his original notes have some anomalous findings in the margins. Next to numbers that do not fit his theory are comments like, “Very low. Something wrong,” and “Possibly a double drop” (Holton 1978, 70). Those oil drops were not included in the final analysis. Did he cheat, and throw away data that didn’t match what he wanted to find? Historian of science Gerald Holton argues that there was probably something observably different about those drops that made Millikan throw out the data points (Holton 1978). In our naïve every-day understanding of science, measurement is a precise process and the things we measure either support or disprove our

hypotheses. In reality, every step of the process can be messy and require active interpretation—struggling to decide whether to exclude possible double drops is the norm rather than the exception.

Science is a messy process, and it is only through verification of results over time that we can judge which findings “are true.” Bruno Latour and Stephen Woolgar document the ways in which the acceptance of a new scientific fact is inherently social (Latour, Woolgar, and Salk 1986). In the field of science studies, this approach is called *constructivism*. At first, in the written scientific literature, it appears as a qualified claim: Millikan and Fletcher claim that the charge on the electron is unitary. As more people duplicate the finding and more people accept it, it might appear with a citation: “since the charge on the electron is unitary (Millikan 1909)....” As it becomes even more broadly accepted, the citation drops out. One can simply refer to the concept without citation because it is broadly accepted as “true.” The need to attribute an idea falls away as the idea is accepted by more people. The process of something being accepted by the scientific community is social.

Of course sometimes the process of creating social consensus makes mistakes. When my father Robert Bruckman first started practicing orthopedic surgery, it was standard practice to apply heat to certain kinds of injuries and cold to others. At some point, the recommendation changed to always use cold. He jokes, “In around the mid 1970s, there was a spontaneous change in the human body, making it more responsive to cold than heat.” Before the mid 70s, it was “true” that applying heat is sometimes a good idea. After the 70s, it was no longer true. Of course the human body didn’t suddenly change—we were wrong about what was best. The process of learning that we are wrong is also social. Anything that we agree “is true” might be wrong. But what we agree is true at any given moment is our best attempt.

Here we get into a philosophical nuance: When someone said heat is good for some injuries, were they “wrong”? Later evidence suggests it. But at the time, we didn’t know. Constructivists like Latour and Woolgar argue that objective truth does not exist independent of a human in a social context who knows that truth at a particular moment in time. Social epistemologists argue that whether we are confused about it or not, objective truth exists (Goldman 1999). I agree with the social epistemologists and Putnam’s internal realism—truth exists but we have only indirect access to it (Lakoff 1987). The distinction is subtle but important. Regardless, for practical purposes, the best we can say is, all knowledge is contingent. Saying something “is true” is linguistic shorthand for, “evidence strongly suggests it, and this is the best we can do until we learn otherwise.”

Peer Review

If truth is socially constructed, how does establishing it work in practice? Peer review is the main mechanism in scientific communities. A finding is reduced to written form as a paper, and that paper is reviewed by experts in the field. The credibility of a paper depends on the rigor of the review process it underwent. Peer review is a way to operationalize social construction of knowledge.

In practice, how carefully something is reviewed depends on the type of publication it has been submitted to. Journal articles are typically the highest standard. Papers submitted to high-quality journals are sent to three experts in the field, and experts review and re-review versions of the paper

until it is approved. Reviewers are typically anonymous, to encourage them to be honest. In some fields, review of conference papers is comparable in rigor to journal articles. In my areas of expertise (human-computer interaction and social computing), conference publications are arguably now more important than journals. In many fields (particularly in the humanities), conference publications are lightly reviewed or not reviewed at all and serve as a way to get feedback from peers on ideas in progress.

We discuss this topic in the undergraduate class I teach on “Computing, Society, and Professionalism,” and it always surprises my students to learn that books are typically less carefully reviewed than other publications. Undergraduate textbooks and books in some humanities disciplines are carefully reviewed, but in other disciplines the review can be cursory. A book manuscript may get comments from a few experts in the field or may not, and the author and publisher are free to ignore comments they don't agree with. People who review book manuscripts may be anonymous, but often are known to the author, which encourages people to err on the side of being diplomatic rather than frank.

Newspaper articles and magazines go through a different form of pre-publication review: fact checking. High quality publication venues have staff assigned to make sure everything they publish is correct. This is both to protect the publication's reputation as reliable, and also to prevent lawsuits for libel. As economic models to support journalism have eroded, many publications no longer fact check or do so less.

Lots of online information receives no formal review. Anyone can put up a blog post or posting on an online discussion site and say whatever they like. Some sites support comments or up and down votes, which provide some feedback on the content. However, those are not equivalent to formal review.

The Reality of Review

It's important not to idealize the process of peer review. Even for a top journal, it's not always what you expect. As research has gotten more inter-disciplinary, it becomes harder for a reviewer to truly understand all the research literature that a new publication is building on. Reviewing is a social process. An editor or conference program committee member requests that others review something. The request typically goes to someone they know. In deciding whether to accept the request, the potential reviewer considers a host of factors including whether they are interested in the topic, whether they understand the area well enough, and whether they owe the requester a favor (did he/she recently review something for me?) Once in the middle of reviewing, the reviewer typically finds parts of the work they understand thoroughly, and other parts that are more unfamiliar. I will include comments for the editor explaining what parts I know better, and what parts I know less well. It's the editor's job to make sure that at least one person is expert on each aspect of the paper, but sometimes that is not how it turns out.

If something I am reviewing cites something I haven't read, should I go read it? In theory I should (and occasionally I do), but it's simply not possible to do that as a matter of standard practice. Given the number of things I'm expected to review and the number of citations each of those contains, I often have to assume that the reference says what the authors say it says. Which is by no means assured.

Ole Bjorn Rekdal (2014) tells the story of the birth of an “academic urban legend”—the idea that spinach is a good dietary source of iron (which is not true). Sloppy failure to check citations perpetuated the myth for decades. Realizing this, a paper by K. Sune Larsson in 1995 wrote that “The myth from the 1930s that spinach is a rich source of iron was due to misleading information in the original publication: a malpositioned decimal point gave a 10-fold overestimate of iron content (Hamblin 1981)” (Larsson 1995). You would think that someone debunking an error in the scientific literature would themselves be careful; however, it turns out that Hamblin had no support for the story of the decimal point and there is no evidence to suggest that really happened (Rekdal 2014). The story about the origin of an academic urban legend is itself an academic urban legend!

Before you read this chapter, did you believe spinach was a good source of iron? Before I read Rekdal’s account, I certainly did. The myth of the iron content of spinach spread and was perpetuated in a community of scientists working in peer-reviewed publications. In this case, the mechanisms to stop such errors (scientific training and peer review) failed. If it’s possible for scientists publishing in journals to create such a persistent falsehood, imagine how easy it is for untrained individuals sharing non-peer-reviewed information on the internet to create falsehoods that are stubbornly hard to kill.

Should You Believe Wikipedia?

As we have seen, the reliability of a publication depends on its degree of review. I will argue that as a result, Wikipedia is sometimes the most reliable publication ever created—and other times not so at all. *It depends on the article.*

Wikipedia can be edited by anyone. Most pages even allow anonymous users to make changes without logging in. Learning that it can be edited by anyone, regardless of their formal qualifications, many people with a traditional bent are repulsed. As we saw in the case of global warming, people are often pre-disposed to believe a particular way. People who for example have a traditional view of authority/credibility are likely to (in the absence of evidence one way or the other) be suspicious of Wikipedia. Those who are inclined to hold positivist, techno-utopian views of the transformative potential of technology are likely (in the absence of evidence) to be excited by it.

The truth is more complicated than it seems on the surface. Consider the case of a really popular article, with many people editing it. Next time someone is nominated to the United States Supreme Court or chosen as the new Pope, take a look at their Wikipedia page. If you click on the “history” tab, you can see every past version of an article, and when it was edited. When someone is suddenly elevated from relative obscurity to a position of prominence, you see Wikipedia at its best.

Consider for example the page of current US Supreme Court Chief Justice John Roberts. His page was created in May 2005, with a few short paragraphs on his educational and professional background and the fact that he served at that time as a judge on the US Court of Appeals for the District of Columbia (182 words total text). The page had nine edits total until July 19th of that year, when he is nominated to the US Supreme Court. That day, it was edited 31 times, and over 1200 times by the end of July. Edits were made in July by 359 different people (174 logged-in users, the rest anonymous users). As of

2020, the text is over 6,800 words long. In addition to reviewing his biography in detail, it reviews positions he has taken as a judge in detail, and links to separate pages about cases he argued before the Supreme Court before becoming a justice, and separate pages about dozens of decisions he made in previous judgeships. His biography is detailed and supported by reliable citations. If you were to insert a random false statement into the article, it would likely be fixed immediately (Viegas, Watternberg, and David 2004). (Please don't do this!)

As we saw before, refereed journal articles are arguably the “gold standard” for quality of information. However, a refereed journal article is reviewed typically by three experts. How do you compare three experts to hundreds of self-selected volunteers? If the volunteers have no real background in the area, they might be able to improve things like readability, grammar and formatting, but might not be able to contribute much to the substance of the article. However, what if they have access to an internet full of information about the topic, and can read references carefully?

Any edit on Wikipedia is typically supported by a reputable citation. If someone adds information without a citation, it is usually quickly removed. When every edit has strong evidence behind it, hundreds of volunteers will do better than a few experts.

Many articles on Wikipedia in fact now do draw communities of experts around them. When my children were young, they suffered from serious food allergies. In the mid 2000s, there was one article on “food allergy.” I created separate articles for some common allergens including egg, dairy, peanut, and tree-nut allergy. In the early days, I was able to contribute to those articles easily—clarifying details or adding new studies. However, over time, the articles have become more specialized with many contributors with professional expertise. The sophistication of Wikipedia editors has increased over time. I still think I could contribute to the articles if I was careful, and experts are there to catch any mistakes I might make. Once you’ve edited an article, it is added to your “watch list,” and Wikipedia editors often check new changes to articles they care about. Experts and non-experts form an intriguing “community of practice” around each article, as we’ll see in Chapter Four.

A journal article goes through the review process once, and then is frozen. If new information emerges to change the consensus view of a phenomenon, subsequent publications can address that, but the original publication doesn’t change (it can only be retracted, if it’s found to be really wrong). A Wikipedia article can be updated on a moment-by-moment basis. In terms of social construction of knowledge, the more people have reviewed something, the more we can trust it. Hundreds of people continually updating a popular Wikipedia page *arguably creates the most reviewed and up-to-date information source ever created*.

While a popular Wikipedia page is heavily reviewed, a less popular Wikipedia page may have hardly been reviewed at all. For example, the page on the Massachusetts State House was created in 2004, and is edited just a handful of times per year. Sections of the article have a “citation needed” tag, warning readers that the section needs additional citations for verification. The page (as of May 2020) includes text saying that:

"Murals on the second floor under the dome were painted by artist [Edward Brodney](#). Brodney won a competition to paint the first mural in a contest sponsored by the [Works Progress Administration](#) in 1936. It is entitled "Columbia Knighting Her World War Disabled." Brodney could not afford to pay models, and friends and family posed. The model for Columbia was Brodney's sister Norma Brodney Cohen, and the model for the soldier on one knee in the foreground was his brother Fred Brodney. In 1938, he painted a second mural under the dome called "World War Mothers." The models were again primarily friends and family members, with sister Norma sitting beside their mother Sarah Brodney. The [New York Times](#) notes that the murals are relatively rare examples of military art with women as their subjects." (Massachusetts State House n.d.)

If you follow the citations, there is support for the fact that Sarah Brodney was the model for the war mothers, but no support for the statement that Norma was the model for Columbia. I promise it's true—I'm the person who added that text in October 2008. Norma was my grandmother, and Sarah was my great grandmother. I would add a cite if I had one. In 2008, you could sometimes get away with adding an un-sourced statement to Wikipedia, if it was on a not particularly popular page. Today, it's less likely. Part of the reason that the mention of Norma has stayed is because *no one cares about the Massachusetts State House*. A low-profile page is much less reliable because it is less carefully reviewed.

Epistemologist Don Fallis notes, "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." The system is self-optimizing: the more people are interested in an article, the more reliable it is likely to be (Fallis 2008).

Other sites often copy text from Wikipedia. A number of other websites now quote the statements about Norma and Fred. Sometimes, this leads to a false citation loop, or "citogenesis" (Frost-Arnold 2018). Here's how this happens. First, someone puts un-sourced content on Wikipedia. Next, someone else writes an article and uses Wikipedia as a source without attribution. Next a Wikipedia editor decides to improve the page by adding better citations, does a web search, finds the article, and adds a citation to it! The appearance of a source is created when in fact there's no support at all. It happens to be true that the models in those murals were Norma and Fred, but if it weren't true, it would be hard to remove that information from the internet. Uneven evaluation of the quality of sources is a weakness of the model.

Separate from the question of whether the contents of an encyclopedia article are true is the question of what it chooses to cover. What topics have articles? Halavais and Lackaff compared the topic of a random sample of Wikipedia articles to representation of those topics in *Books in Print*, and found that it varies—for example the sciences are better represented than the social sciences. Comparing Wikipedia to specialized dictionaries, they found that Wikipedia's coverage of physics and linguistics are better than its coverage of poetry. And of course, topics of general interest like popular culture are disproportionately well covered—there's a whole lot that's been written about Harry Potter and Star Wars (Halavais and Lackaff 2008). Wikipedia's coverage continuously expands, so how well it represents any area is constantly changing (Mesgari et al. 2015).

If an article exists on a topic, how thorough is it? What things are included and what things are left out? Historian Roy Rosenzweig (2006) compared 25 biographies on Wikipedia to the same subjects covered in a scholarly work, American National Biography Online (ANBO). Both were reliable—Rosenzweig found four minor, factual errors in those 25 articles in Wikipedia and one in ANBO. More interesting, though, is differences in the style of the articles and what they choose to include. Comparing the biographies of American President Abraham Lincoln, both are accurate. ANBO has elegant prose, richer contextualization, connections to themes in historiography, and confident judgment. Wikipedia has an explicit policy demanding that articles have a “neutral point of view” (NPOV), so judgment is not allowed. The Wikipedia article was also written by many people, and as a result can’t have such flowing prose. However, Wikipedia has many fun, colorful details about Lincoln that are not present in ANBO (like the fact that he shared a birthday with Charles Darwin, or that he wrote the edict to make Thanksgiving a national holiday in the United States) (Rosenzweig 2006).

In “Towards an Epistemology of Wikipedia, Don Fallis (2008) concludes that Wikipedia is largely reliable, but also reminds us that there are other epistemic virtues beyond reliability—notably “power, speed, and fecundity. 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” (Fallis 2008). Wikipedia excels by those measures.

Theoretical Summary

It’s surprising how much philosophy is needed to answer a simple question—whether you should believe Wikipedia. The underlying philosophical questions like whether objective truth exists have no simple answers. My own thinking tends to embrace common sense, moderate compromises. The view I have cobbled together for myself is, in sum:

The world exists, but is only knowable through our fallible senses (internal realism). Truth likely exists independent of knowledge and the knower, but we have only human ways of attempting to gain access to it. Because our senses are fallible, the way we determine what we believe “is true” is through social consensus (social epistemology). Social consensus is sometimes wrong, and we are constantly improving on it.

Knowledge requires true, justified belief. Knowledge is an achievement of a reliable cognitive process. Sharing knowledge is a collaborative achievement, and key features of the social environment shape whether that process is likely to be successful.

Wherever you personally land on these thorny questions, insights from epistemology (what is true) and sociology of knowledge (what people choose to believe) are becoming surprisingly useful on a day-to-day basis. Not just Wikipedia but the entire internet is a swirling ecosystem of social construction of knowledge in action.

Practical Implications

Knowing something about the nature of knowledge and truth, what can we do better as users of the internet and designers of internet-based communication systems? At the simplest level, it helps to be aware of the fragility of all knowledge. Epistemologists can debate what a “basic fact” is and what it means for something to be really supported, but on a day-to-day basis none of us have time to dig down into support for ideas until we hit some kind of basic fact bedrock (if such a thing even exists). We are not all climate scientists—if we want to decide whether it’s worth extra money to buy a low emissions car, we need to choose sources we trust. And no one can be knowledgeable about everything—even actual climate scientists need to find sources they trust about whether to vaccinate their kids, eat gluten, or buy produce from genetically modified crops.

Understanding more about social construction and the nature of knowledge, we can be better consumers of information and better designers of information systems. Consider this example: Recently I accidentally retweeted a false news story. The story stated that drivers for ride sharing services Uber and Lyft were making a median profit of \$3.37 per hour—shockingly low (Levin 2018). The article was posted by the newspaper *The Guardian* (arguably one of the most reputable news sources in the world), and was based on a study done at MIT (arguably one of the most prestigious research institutions in the world). With that pedigree, I retweeted the story without thinking about it too deeply. Friends on Twitter immediately expressed skepticism—that number can’t be right. Why would so many people be driving for those services if the pay was so low? Digging deeper, it turns out that the MIT paper was just a working paper—it hadn’t undergone peer review. In fact, after criticism from Uber, the MIT authors posted a correction—the median is \$8.55 per hour (a much more reasonable figure). I quickly deleted my posting, and tweeted a correction. By posting something false without realizing it, I contributed to the problem of *misinformation*. Sharing something false deliberately is *disinformation*, and I’ll talk about that in Chapter Six.

There’s bad news and good news in this story. The bad news is that sometimes you can’t even trust some of our most hallowed brands, *The Guardian* and MIT. If those information sources aren’t always credible, how are we to navigate the sea of articles by much less reputable sources? But the good news is, in this case the system worked. My friends immediately expressed skepticism, and I took down my tweet and posted a correction. The study authors checked their figures and posted a correction. *The Guardian* posted a correction. Viewing this through the lens of an ecology of social construction of knowledge, we can look at what specific design features led to the original error, and what features helped generate a correction.

How do we support people in making well-grounded decisions about what to believe and what to share? One possibility is to provide people with meta-data: information on the sources of support for a given piece of information (like the fact that the *Guardian* story was based on a non-peer-reviewed working paper). Many sites have begun efforts to improve in this area. For example, in 2017, Facebook started adding more context to articles (Welch 2018). This includes a link on articles with a small “i” that takes you to a page with the date the article was first shared and a profile of the publisher. This source information is only available on a small fraction of articles, and mainly on articles from major news sources whose reputations are already widely known. It’s unfortunately not available where it’s needed most—on articles from smaller sites whose reputations may be unclear. In February 2020,

Twitter announced that they would start labeling synthetic and manipulated media (Roth and Achuthan 2020). In May 2020, they announced that they would start labeling “potentially harmful and misleading content,” particularly related to the coronavirus pandemic (Roth and Pickles 2020), and in late May for the first time marked two tweets by the President of the United States as misleading (Wong and Levine 2020). More and more of his tweets were marked as misinformation over time, until he was finally banned from the platform in January 2021.

Can meta-data help? This question leads to a long list of harder questions. Who decides what tweets or articles to mark as misleading or false? By what standard? Assuming the annotation is done well, does marking things as misleading or false help the reader? What design features of the user interface for meta-data make it more effective? When does meta-data backfire, leading people to be more convinced of the truth of the false claim (Nyhan and Reifler 2016)?

What kinds of meta-information are helpful? How detailed should that information be? How do we generate meta-data? Do we use human labor or artificial intelligence to generate it? How do we train the human workers and the algorithms? Do we need multiple sources of meta-data, with different underlying values and assumptions? Is there a potential demand for a “reliability service”? Is there a potential business model to support such a service? What approaches other than meta-data might help us address the problem of false and misleading information on the internet?

The people struggling with these questions are diverse, including teachers, students, researchers, publishers, and citizens trying to make every possible kind of personal and political decision. We can all make better decisions about what to believe if we are armed with a deeper understanding of the nature of “truth” and “knowledge.”

In the next chapter we’ll see that part of the problem is who is sharing information with whom—if you don’t believe the moon landings actually happened, you are more likely to persist in that belief if you are part of a group who all agree with you (Sunstein 2018).

Bibliography

- Borges, Jorge Luis, and Andrew Hurley. 1999. *Collected Fictions*. Penguin Books.
- Fallis, Don. 2008. "Towards an Epistemology of Wikipedia." *Journal of the American Society for Information Science and Technology* 59 (10): 1662-1674.
<https://doi.org/https://doi.org/10.1002/asi.20870>.
- Feldman, Richard. 2003. *Epistemology*. NJ: Prentice Hall.
- Frost-Arnold, Karen. 2018. "Wikipedia." In *The Routledge Handbook of Applied Epistemology*, edited by David Coady and James Chase. Routledge.
- Goffman, Erving. 1959. *The Presentation of Self in Everyday Life*. New York: Doubleday.
- Goldman, Alvin I. 1999. *Knowledge in a Social World*. Oxford University Press.
- Greco, John. 2021. *The Transmission of Knowledge*. Cambridge University Press.
- Halavais, Alexander, and Derek Lackaff. 2008. "An analysis of topical coverage of Wikipedia." *Journal of Computer-Mediated Communication* 13 (2): 429-440.
<https://doi.org/https://doi.org/10.1111/j.1083-6101.2008.00403.x>.
- Hamblin, T. J. 1981. "Fake!" *British Medical Journal* 283 (6307): 1671.
<https://doi.org/https://dx.doi.org/10.1136%2Fbmj.283.6307.1671>.
- Heersmink, Richard. 2017. "A Virtue Epistemology of the Internet: Search Engines, Intellectual Virtues, and Education." *Social Epistemology* 32 (1): 1-12.
<https://doi.org/10.1080/02691728.2017.1383530>.
- Holton, Gerald. 1978. *The Scientific Imagination: Case Studies*. Cambridge University Press.
- An Inconvenient Truth. n.d. *Wikipedia*.
https://en.wikipedia.org/wiki/An_Inconvenient_Truth.
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things; What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.
- Larsson, K. Sune. 1995. "The dissemination of false data through inadequate citation." *Journal of Internal Medicine* 238 (5): 445-450.
<https://doi.org/https://doi.org/10.1111/j.1365-2796.1995.tb01222.x>.
- Latour, Bruno, Steve Woolgar, and Jonas Salk. 1986. *Laboratory Life*. Princeton, NJ: Princeton University Press.
- Levin, S. 2018. "Uber drivers often make below minimum wage, study finds." *The Guardian*, March 5th, 2018. .
<https://www.theguardian.com/technology/2018/mar/01/uber-lyft-driver-wages-median-report>.
- Massachusetts State House. n.d. *Wikipedia*.
https://en.wikipedia.org/wiki/Massachusetts_State_House#Inside_the_building.
- Mesgari, Mostafa, Chitu Okoli, Mohamad Mehdi, Finn Arup Nielsen, and Arto Lanamaki. 2015. "'The sum of all human knowledge': A systematic review of scholarly research on the content of Wikipedia." *Journal for the Association for Information Science and Technology* 66 (2): 219-245.
<https://doi.org/https://doi.org/10.1002/asi.23172>.
- Millikan, Robert Andrews. 1909. "A new modification of the cloud method of measuring the elementary electrical charge, and the most probable value of that charge." *Physical Review* 29 (6): 560-561.
<https://resolver.caltech.edu/CaltechAUTHORS:MILpr09>.
- Nickerson, Raymond S. 1998. "Confirmation bias: A ubiquitous phenomenon in many guises." *Review of General Psychology* 2 (2): 175-220.
<https://doi.org/https://doi.org/10.1037%2F1089-2680.2.2.175>.
- Nyhan, Brendan, and Jason Reifler. 2016. "When Corrections Fail : The Persistence of Political Misperceptions " *Political Behavior* 32 (2): 303-330.
<http://www.jstor.org/stable/40587320>.
- Rekdal, Ole Bjorn. 2014. "Academic urban legends." *Social Studies of Science* 44 (4): 638-654. <https://doi.org/https://doi.org/10.1177%2F0306312714535679>.
- Rosenzweig, Roy. 2006. "Can history be open source? Wikipedia and the future of the past." *Journal of American History* 93 (1): 117-146.
<https://doi.org/10.2307/4486062>.

- Roth, Yoel, and Ashita Achuthan. 2020. Building rules in public: Our approach to synthetic & manipulated media.
https://blog.twitter.com/en_us/topics/company/2020/new-approach-to-synthetic-and-manipulated-media.html.
- Roth, Yoel, and Nick Pickles. 2020. Updating our approach to Misleading Information. https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.html.
- Steup, Matthias. 2016. Epistemology. *Stanford Encyclopedia of Philosophy*.
<https://plato.stanford.edu/archives/fall2016/entries/epistemology/>.
- Sunstein, Cass R. 2018. #Republic : divided democracy in the age of social media. Princeton University Press.
- Viegas, Fernanda, Martin Wattenberg, and Kushal David. 2004. "Studying cooperation and conflict between users with history flow visualizations." ACM CHI, Vienna, Austria, April. <https://doi.org/10.1145/985692.985765>
- Welch, C. 2018. Facebook will give all US users helpful context about articles shared in News Feed. *The Verge*.
<https://www.theverge.com/2018/4/3/17193750/facebook-news-feed-articles-publisher>.
- Wong, Carrie Julia, and Sam Levine. 2020. "Twitter labels Trump's false claims with warning for first time." *The Guardian*, May 26, 2020.
<https://www.theguardian.com/us-news/2020/may/26/trump-twitter-fact-check-warning-label>.
- Word of the Year. 2016. *Oxford Dictionaries*.
<https://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016>.