

Can You Trust Dr. Wikipedia? Studying the accuracy of Wikipedia health science articles is not easy.

By Jonathan Jarry M.Sc. | 6 Sep 2024 Office for Science and Society, Separating Sense from Nonsense
McGill University, Montreal, Canada

Do you know who invented the electric toaster? If you answered Alan MacMasters, a young Scotsman with high cheekbones and quite a head of hair, you've been lied to by Wikipedia.

That infamous prank was pulled by two university students after one of their professors told them not to use Wikipedia for research purposes. "Because you never know," he warned, "who might set themselves as the inventor of the toaster." So they did. The hoax, which was repeated as fact in major media outlets like the BBC, lasted nearly a decade.

There's also the time when the assistant to Robert Kennedy, John Seigenthaler Sr., was mentioned on Wikipedia as having been briefly suspected of a direct involvement in the assassinations of both Robert and his brother, John F. Kennedy. It turns out that a Tennessee man had tried to trick a co-worker by adding the false information on Wikipedia.

No wonder so many professors are reputed for ushering their students away from Wikipedia as a reliable reference.

But Wikipedia, which is written and edited by volunteers and has existed since 2001, is surprisingly accurate. It's difficult to answer a question as broad as "is Wikipedia reliable?" given the 63 million articles it offers, with roughly one-tenth being in English. Bearing in mind our Office's focus on science and my own background in the health sciences, I will put aside other subjects and plunge into some of the studies that have been done on Wikipedia's scientific accuracy. It turns out it's not an easy thing to measure.

Having your cake and eating it too

The first major study of Wikipedia's accuracy on scientific issues was published in the prestigious *Nature* academic journal a mere four years after Wikipedia's launch. It pitted the budding website against the undefeated champion, *Encyclopedia Britannica*. The take-home message? Wikipedia "came close" to the encyclopedia in terms of being accurate. The editors of *Britannica* disputed the findings; *Nature* pushed back.

I have seen this study cited in nearly every scientific paper analyzing the reliability of Wikipedia, and it is historically important. But Wikipedia has grown tremendously both in terms of quality and quantity since 2005. This is a problem with research into this topic: we are getting snapshots in time that themselves quickly become unreliable. One academic wrote in 2008 that "the site's volatility diminishes its credibility." That same volatility also taints the work of researchers trying to study it.

Another problem is sampling. Given the vastness of Wikipedia's content, it is impossible for a research team to examine every single published entry even when limiting itself to a specialty. Researchers have to pick a sample and hope it is representative. *Nature* investigation, for example, only looked at 42 articles. When we survey studies done on drug information accuracy on Wikipedia, the results are all over the map. A 2023 investigation into 100 chemotherapeutic drugs chosen at random reported that nearly two-thirds of these articles were of poor quality, missing basic information about routes of administration and contraindications. But a study done nine years earlier on another set of 100 drugs also chosen at random (though not necessarily used for chemotherapy) came to a different conclusion: accuracy was stellar and completeness was quite good compared to pharmacology textbooks. Are we to conclude that Wikipedia has worsened significantly in the last decade? No. Rather, these studies show the limitations of sampling.

But it's not just sampling; there's also the question of what researchers want Wikipedia to be. Especially on medical topics, many academics evaluating the reliability of Wikipedia want to have it both ways, it seems. They at once denounce errors of omission *and* the complexity of the language used. They compare Wikipedia to textbooks aimed at healthcare professionals *and* they lament the fact that the Wikipedia articles are too dense to be understandable by the average person. I want to eat a whole pie without gaining weight, feeling sick, or seeing a spike in my blood sugar, but I realize that we can't always get what we want. A single Wikipedia article cannot be as thorough as a medical textbook yet as approachable as a patient leaflet. (Wikipedia's answer to the problem: Simple English Wikipedia.)

The literature appraising Wikipedia articles is also heavily biased toward the English-language version of the encyclopedia (and, to be fair, so were my own

searches for these papers). But an English-language Wikipedia page and its French equivalent are not identical. A study on how the eye disease known as diabetic retinopathy was covered by Wikipedia in 19 different languages showed that quality varied immensely, from “fair” for the English and German versions to “poor” for most of the languages consulted. And because Wikipedia is decentralized, sources that have been deemed untrustworthy by English Wikipedians are still cited in Wikipedia pages produced in other languages.

Looking at Wikipedia through a microscope is thus really hard, but there are trends that emerge from these imperfect snapshots. Wikipedia, overall, has no business being this good. Given the deterioration so often witnessed on social media platforms (and a friend’s horrifying, four-year experiment in running an unmoderated free speech group on Facebook), I originally expected Wikipedia to devolve into a cesspool, with articles vandalized by trolls and offering all sorts of dangerous misinformation. But that’s not the case. The site is far from perfect, and it will not replace dedicated professional databases for advanced knowledge (nor should it), but as a point of entry into a topic, it is quite reliable overall.

And there are good reasons for that.

Turning students into editors

Despite Wikipedia’s free-for-all appearance, not everyone can edit or write an article on the site. Only users registered on Wikipedia can create new articles, which sets up a small barrier to entry aimed at dissuading spur-of-the-moment pranksters. Controversial topics can also be protected. The page for Donald Trump, for example, has a little blue lock icon on it with the letter “E” inside of it, indicating an “extended confirmed protection.” Only users whose accounts are at least a month old and who have performed a minimum of 500 edits on Wikipedia are allowed to edit this page. Moreover, dedicated Wikipedians keep an eye on trending Wikipedia pages (which might indicate that a topic is being hotly debated on social media) and ensure that no lies or vandalism stick around for long.

Since at least 2013, health science pages on Wikipedia have additionally benefitted from the organized participation of a particular subset of people: students. Some universities are offering elective courses that train health science students on how to make edits on Wikipedia. These students are then unleashed on important pages that have failed to meet the standards of the online encyclopedia. At the University of Notre Dame Australia, every first-year medical student learns how to edit Wikipedia during their orientation week. Some of these isolated initiatives belong to a larger affinity group called WikiProject Medicine, which played an important role in curbing misinformation on Wikipedia during the early days of the COVID-19 pandemic.

Even on pseudoscientific topics, which invite controversy and strong emotions, the encyclopedia is surprisingly good. This is in part due to the work done by Guerilla Skepticism on Wikipedia, ensuring that these pages fall on the side of the evidence and not wishful thinking. (In the interest of full disclosure, their off-Wikipedia organizing has been the subject of criticism.)

Wikipedia is not perfect; but as has been argued before, denying students the use of the encyclopedia does not teach them anything. There's a reason why so many students gravitate toward it. When 116 medical students in Canada took a short test similar to their licensing examination, they were asked to write down the topics they had had difficulty with afterwards. They were then randomized to looking those topics up either in a medical textbook, on the website UpToDate (commonly used by doctors), or on Wikipedia. They then took the test again. Who did best? The students who brushed up on their knowledge on UpToDate... or on Wikipedia. The latter was praised by the students for being easy to navigate and for having practical hyperlinks. It was both useful and fast.

After all, the word "wiki" is Hawaiian for "quick."

1. Qual o significado da frase "Separating Sense from Nonsense"? (1,0)
 - a. Separando sentido de não sentido.
 - b. Separando bom senso de sandices.
 - c. Separando significados de não significados.
 - d. Separando senso de não senso.
2. No parágrafo 2, qual a razão dada pelo professor para não usar Wikipedia? (1,0)
 - a. Qualquer um pode se tornar autor da Wikipedia.
 - b. Qualquer um pode opinar sobre qualquer assunto.
 - c. Qualquer um pode se colocar como inventor da torradeira.
 - d. Qualquer um pode descrever o uso correto da torradeira.
3. O que os professores geralmente fazem segundo o parágrafo 4? (1,0)
 - a. Os professores guiam os alunos para usos corretos do site.
 - b. Os professores encorajam a comparação de informação.
 - c. Os professores encorajam o uso da *Encyclopedia Britannica*.
 - d. Os professores desencorajam o uso da Wikipedia.

4. Quais as dificuldades em analisar a Wikipedia em termos da confiabilidade?
Parágrafo 5. Escolha a opção incorreta. (1,0)
- a. Há 63 milhões de artigos na Wikipedia .
 - b. Os artigos são escritos por voluntários de vários países.
 - c. Os artigos são editados por profissionais contratados.
 - d. Os artigos são escritos em muitas línguas.
5. Qual o significado do subtítulo, “Having your cake and eating it too”? (1,0)
- a. Guardando o pedaço de bolo e comendo o pedaço.
 - b. aproveitar as partes boas de algo sem ter ou lidar com as partes ruins.
 - c. Uma prática de guardar um pedaço do bolo de casamento.
 - d. Um ditado que significa, É preciso pensar no amanhã.
6. Qual o significado de “The take-home message”? (1,0)
- a. O que se aprende de algo.
 - b. A mensagem que é explicada.
 - c. O que se carrega para o futuro.
 - d. A mensagem que é ouvida.
7. Segundo o texto, quais os problemas de estudos de amostragem? Escolha a opção incorreta. (1,0)
- a. Há inúmeros assuntos abordados na Wikipedia que dificultam os estudos.
 - b. Os estudos de amostragem sempre têm resultados idênticos.
 - c. Ao escolher as amostras, não há a certeza de que serão representativas.
 - d. Haverá resultados diferentes em cada estudo de amostragem.
8. Por que Wikipedia criou “Simple English Wikipedia”? Escolha a opção correta. (1,0)
- a. Para ajudar o leitor de inglês como segunda língua.
 - b. Para que os escritores dos artigos tivessem mais facilidade de escrita.
 - c. Para ajudar o leitor leigo a compreender assuntos complexos.
 - d. Para ser mais próximo aos artigos da *Encyclopedia Britannica*.
9. Como a Wikipedia se protege? Escolha a opção incorreta. (1,0)
- a. É preciso se registrar na Wikipedia para poder escrever artigos.
 - b. Assuntos controversas são protegidos e fechados à edição aleatória.
 - c. Os próprios autores dedicados à Wikipedia vigiam as novas páginas.

d. É preciso provar a competência numa área específica para ser autor.

10. Por que os alunos de medicina mencionados no artigo gostaram da Wikipedia?

Escolha a opção incorreta. (1,0)

a. Porque é útil e rápido.

b. Porque o nome significa rápido em Havaiano.

c. Porque tem hipertextos.

d. Porque é fácil de navegar.