The Professional Geographer

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rtpg20

A Preliminary Investigation of Fake Peer-Reviewed Citations and References Generated by ChatGPT

Terence Day

To cite this article: Terence Day (2023): A Preliminary Investigation of Fake Peer-Reviewed Citations and References Generated by ChatGPT, The Professional Geographer, DOI: 10.1080/00330124.2023.2190373

To link to this article: https://doi.org/10.1080/00330124.2023.2190373

Published online: 12 Apr 2023.

A Preliminary Investigation of Fake Peer-Reviewed Citations and References Generated by ChatGPT

Terence Day
Okanagan College, Canada

An analysis of academic citations and references generated by the ChatGPT artificial intelligence (AI) chatbot reveals the citations and references are in fact, fake. They are clearly generated by a predictive process rather than known facts. This suggests that early optimism regarding this technology for assisting in research could be misplaced, and that student misuse of the chatbot can be detected by the identification of fake citations and references. Despite these problems, the technology could have application in the writing of course materials for lower level undergraduate courses that do not necessarily require references. Subject matter expertise is required, however, to identify and remove incorrect information. The need to identify incorrect information provided by an AI chatbot is a skill that students will also increasingly need.
Key Words: AI, artificial intelligence, chatbot, ChatGPT, education.

ChatGPT is an artificial intelligence (AI) chatbot that uses natural language to answer questions. It has quickly become a cultural sensation (Thorp 2023) since it trial launched in November 2022, with, "far-reaching consequences for science and society" (van Dis et al. 2023), and with "remarkable performance" on medical licensing exams (Kung et al. 2023). Concerns about the impact of the technology on academic research, and on student academic integrity also quickly emerged, however (Rudolph, Tan, and Tan 2023).

In February 2023 I used the ChatGPT February 13 version (see https://chat.openai.com/) to investigate a few questions in areas of physical geography and geography education in which I have some knowledge. A number of results generated unfamiliar citations and references. On checking, the references were found not to exist in the journal to which they were attributed, nor anywhere else.

This research note documents the results of a subsequent more systematic but preliminary investigation of the veracity of citations and references generated by the ChatGPT chatbot.

Method

As a test of the veracity of references generated by ChatGPT, a systematic review was undertaken. I posed questions to ChatGPT inspired by five recent papers in *The Professional Geographer*. The questions (Table 1) covered participation of women, minorities, and people with disabilities in K–12 geography (Solem 2023), reasons for migration of people to the U.S. West (Otterstrom and Shumway 2023), the geography of population growth in China (Shi et al. 2023), use of digital maps (Bravo and Sluter 2023), and the role of imagination in environmental management (Chhetri, Ghimire, and Eisenhauer 2023). There was no specific reason for choosing these questions; they are simply a qualitatively representative sample of the sorts of questions commonly asked by geographers.

In all cases the references supplied by ChatGPT were checked on the Web sites of the cited journal, with an additional check on Google Scholar based on the title in quotation marks to verify that there had not been a simple typographic error in the journal name, volume, issue, or page number.

Web citations were entered into a browser and checked on the Internet Archive "Wayback Machine" (see https://web.archive.org/).

Results

In no cases could the references be confirmed to be genuine (Table 1).

The Li and Montello paper is related to, "Montello, D. R., Waller, D., Hegarty, M., & Richardson, A. E. (2004). Spatial memory of real environments, virtual environments, and maps. In Human spatial memory (pp. 271–306). Psychology Press." However, the only similarity is the title and the inclusion of Montello as an author.

The article by Epstein et al. is identified as possibly referring to the widely cited paper, "Epstein, G., Pittman, J., Alexander, S. M., Berdej, S., Dyck, T., Kreitmair, U., … & Armitage, D. (2015). Institutional fit and the sustainability of social–ecological systems. *Current Opinion in Environmental Sustainability*, *14*, 34–40." In the ChatGPT version of this reference, the authors are changed slightly (Krebs instead of Kretmeir), the source is incorrectly stated as "Ecology and Society," and the volume and page numbers are assigned in the journal to

The Professional Geographer, 0(0) 2023, pages 1–4
Initial submission, March 2023; final acceptance, March 2023.
Published by Taylor & Francis Group, LLC.

Table 1 *The questions and the references*

Question	ChatGPT supplied references	Fake?
Citing peer-reviewed sources with references, is K–12 geography open to the participation of women, minorities, and persons with disabilities?	Boehm, R. G., & Sparks, B. G. (2019). We have a diversity problem: Minority and women's underrepresentation in geography higher education. *Journal of Geography in Higher Education, 43*(3), 308–322.	Fake
	Goodwin, M., & Koschmann, M. (2014). Accessible geography: Removing the barriers for students with disabilities. *Journal of Geography, 113*(2), 66–73.	Fake
Citing peer-reviewed sources, are people drawn to the American West for its resources, or for aesthetic reasons?	Fiege, M. (2010). The West as America's story: The myth of the West in environmental history. *Environmental History, 15*(2), 195–212.	Fake
	Gunn, S., & Kellner, K. (2014). Landscape and identity in the American West. *Journal of Cultural Geography, 31*(1), 1–23.	Fake
	Williams, M. L., & Murdoch, J. (2012). Natural resource use and economic growth in the American West. *Annals of the Association of American Geographers, 102*(4), 816–834.	Fake
Citing peer-reviewed sources, have there been recent changes in the geography of population growth in China?	Cai, Y., & Wang, F. (2013). China's demographic and spatial-economic transformations: Recent trends and future challenges. *Population and Development Review, 38*(s1), 271–292.	Fake
	Fan, J., & Chen, J. (2013). Population change in China: Urbanization and decentralization. *Journal of Geographical Sciences, 23*(1), 7–22.	Fake
	Li, X., & Hu, Z. (2016). Population growth and land use change in China's western regions: A case study of Chongqing. Land Use Policy, 58, 224-231.	Fake
Do digital maps change the way that map-use tests need to be conducted?	Arend, B. D., & Hwang, G. (2015). Comparing paper and digital maps: Insights from cognitive science. *Cartography and Geographic Information Science, 42*(sup1), 3–13.	Fake
	Golledge, R. G. (1999). Wayfinding behavior and cognitive mapping: Some reflections on the state of the art. *Journal of Environmental Psychology, 19*(1), 33–40.	Fake
	Li, R., & Montello, D. R. (2010). Spatial memory of real environments, virtual environments, and maps. *Journal of Environmental Psychology, 30*(4), 474–482.	Fake, but see comments in text
	Liu, X., Liao, H., & Fan, H. (2018). Spatial ability requirements for using paper maps and digital maps: Evidence from a map-based route learning task. *International Journal of Human-Computer Studies, 114,* 1–9.	Fake
Does effective environmental management benefit from imagination?	Epstein, G., Pittman, J., Alexander, S. M., Berdej, S., Dyck, T., Krebs, J., … & Villamayor-Tomas, S. (2015). Institutional fit and the sustainability of social-ecological systems. *Ecology and Society, 20*(3), 45.	Apparently genuine, but incorrect citation
	Ingram, J. C., & Franco, C. (2014). Imaginative engagement with environmental issues: Exploring the ecolinguistic dimensions of environmental narratives. *Environmental Science & Policy, 37,* 23–33.	Fake
	Levine, A., & O'Neill, K. (2014). Environmental imagination and environmental action: Imaginative agency and the prospects for a sustainable future. *Sustainability Science, 9*(4), 437–451.	Fake
	Wals, A. E. J., & van der Hoeven, N. (2011). Learning for sustainability and the imagined future. *Journal of Cleaner Production, 19*(9–10), 995–1001.	Fake

completely different papers. According to the ChatGPT answer, the Epstein et al. study "suggests that imagination can be a powerful tool for environmental management, particularly in situations where there is uncertainty or complexity in the ecological or social systems being managed. The authors argue that imagination can help managers to envision and evaluate alternative scenarios or future states, and can help to identify and address potential risks or unintended consequences of management actions." A reading of the real paper, however, shows it does not mention the role of imagination in environmental management, and therefore the ChatGPT summary of this incorrectly cited paper is also fake.

Two Web citations were also provided by ChatGPT for the first question, namely https://www.ncge.org/publications/guidelines-for-geographic-education-equity-social-justice-and-sustainability and https://www.nationalgeographic.org/education/geo-inquiry/geo-inquiry-process/. Neither of the URLs worked, and there was no trace of them on the Internet Archive Wayback Machine.

Discussion

Although the reference and citations are in English, in highly regarded journals, and look legitimate on

first examination, they are all fake and hard to detect on initial examination. The citations are in a consistent style and the page numbers fall within the volume and issue of the real journal, but in all cases the relevant journal pages are occupied by papers unrelated to the listed reference.

Many of the authors listed by ChatGPT are highly regarded in their field, but they did not write the papers they have been credited with. The deceptive power is reinforced by the fact that the referenced journals are real and publish genuine articles related to each theme. For example, the *Journal of Geography* and the *Journal of Geography in Higher Education* focus on geography education, and it is plausible, and perhaps expected, that questions relating to geography education would reference papers from these journals. Similarly, for the question on China, genuine relevant papers have been published in *Population and Development Review* and *Land Use Policy* (e.g., Wang et al. 2017; Xiong 2022; Yin et al. 2022; Wen et al. 2023). The other cited journal, *The Journal of Geographical Sciences*, has not only published relevant research; it is also sponsored by the Geographical Society of China. These are all credible and plausible outlets for relevant research.

The use of ChatGPT in the research process must be seriously questioned. Any acknowledgment to ChatGPT or even ChatGPT "coauthorship" should alert editors to potential problems with the manuscript. The presence of multiple incorrect references and no acknowledgment of ChatGPT could signal malfeasance on the part of authors. The peer review and editorial process is now complicated by the need to check not just the format of citations and references, but also their existence and accuracy, which casts doubt on an early suggestion that "Chatbots provide opportunities to complete tasks quickly, from PhD students striving to finalize their dissertation to researchers needing a quick literature review for their grant proposal, or peer-reviewers under time pressure to submit their analysis" (van Dis et al. 2023). The results might look plausible, but they fail any reasonable test.

From the perspective of student academic integrity, the presence of fake references and citations provides a useful tool to check for violations of university or college policies or codes of academic integrity. It also imposes an additional burden on anyone who grades student academic work, however. This burden could be alleviated by requiring students to provide a URL to the digital object identifier (DOI) of their references. If the paper fails to appear when the link is clicked, the paper is probably fake, or at least requires additional investigation.

Despite the problems enumerated, the titles of the fake articles are all directly relevant to the questions and could potentially make excellent papers. The lack of a genuine citation could signal an opportunity for an enterprising author to fill a void.

There could also be other applications. Professors and lecturers working as content curators could also potentially use the AI chatbot to assist in rapid writing of introductory course materials that do not require referencing. For example, I am starting to use it for writing course materials for student use in introductory physical geography courses, for my second-year course in geographical hydrology, and second-year cartography, geographic information systems, and remote sensing. ChatGPT does not produce maps or diagrams, so they still need to be produced or found elsewhere. I have found, though, that it is possible to produce 10,000 to 15,000 words of high-quality written material for students in a day or two. This approach could have significant relevance for open learning and could potentially affect current textbook publishing models.

There are limitations to the use of ChatGPT for the writing of course materials. I have experimented with it for my third-year course in the physical geography of emerging economies, but with much less success. The course relies too much on primary research papers, and in addition to fake papers, ChatGPT generated dubious subject content. Even in that situation, though, AI chatbot answers did serve as a prompt or foil for ideas.

I have also used incorrect information from ChatGPT as the basis for midterm test questions in my second-year hydrology and geomatics courses. Students were required to identify the error(s) in paragraphs. Subsequent informal discussions with students suggest they immediately understood the relevance of this type of question to their future.

Conclusions

It is clear that references generated by ChatGPT are fake and an artifact of the process used to generate the answers. Although the work was undertaken in the context of academic geography, the problem is caused by the AI algorithm, not by the subject matter. Therefore, the results have broad applicability and apply to the use of ChatGPT in all disciplines.

The discovery of fake sources also calls into question the veracity of statements made in ChatGPT answers to questions. The chatbot will likely improve, but initial enthusiasm should be tempered with a more nuanced and cautious approach to the application of AI chatbot technology to teaching and research.

It is unlikely the technology will disappear because of the teething problems identified here and there are still many potential applications of AI chatbots in postsecondary education. It is a tool that requires subject expertise, however. ■

Acknowledgment

I appreciate helpful conversations with Marie Molloy and Gill Green during the writing process. Two anonymous reviewers made constructive comments and suggestions. Thank you.

ORCID

Terence Day http://orcid.org/0000-0001-7140-6794

Literature Cited

Bravo, J. V. M., and C. R. Sluter. 2023. Crowdsourcing map-using and map-generating tasks into OpenStreetMap. *The Professional Geographer* 75 (2):248–62. doi: 10.1080/00330124.2022.2094424.

Chhetri, N., R. Ghimire, and D. C. Eisenhauer. 2023. Geographies of imaginaries and environmental governance. *The Professional Geographer* 75 (2):263–68. doi: 10.1080/00330124.2022.2087698.

Kung, T. H., M. Cheatham, A. Medenilla, C. Sillos, L. De Leon, C. Elepaño, M. Madriaga, R. Aggabao, G. Diaz-Candido, J. Maningo, et al. 2023. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health* 2 (2):e0000198. doi: 10.1371/journal.pdig.0000198.

Otterstrom, S. M., and J. M. Shumway. 2023. New west and old west in the twenty-first century: The rich get richer. *The Professional Geographer* 75 (2):220–32. doi: 10.1080/00330124.2022.2103719.

Rudolph, J., S. Tan, and S. Tan. 2023. ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching* 6 (1):1–22. doi: 10.37074/jalt.2023.6.1.9.

Shi, Q., L. Tao, Z. Yunxia, and P. Rongxi. 2023. People and places in the 2020 census: New geographies of population growth in China? *The Professional Geographer* 75 (2):233–47. doi: 10.1080/00330124.2022.2103721.

Solem, M. 2023. Geography achievement and future geographers. *The Professional Geographer* 75 (2):207–19. doi: 10.1080/00330124.2022.2081227.

Thorp, H. H. 2023. ChatGPT is fun, but not an author. *Science* 379 (6630):313. doi: 10.1126/science.adg7879.

van Dis, E. A., J. Bollen, W. Zuidema, R. van Rooij, and C. L. Bockting. 2023. ChatGPT: Five priorities for research. *Nature* 614 (7947):224–26. doi: 10.1038/d41586-023-00288-7.

Wang, X., Y. Bai, L. Zhang, and S. Rozelle. 2017. Migration, schooling choice, and student outcomes in China. *Population and Development Review* 43 (4):625–43. doi: 10.1111/padr.12101.

Wen, Q., J. Li, J. Ding, and J. Wang. 2023. Evolutionary process and mechanism of population hollowing out in rural villages in the farming-pastoral ecotone of northern China: A case study of Yanchi County, Ningxia. *Land Use Policy* 125:106506. doi: 10.1016/j.landusepol.2022.106506.

Xiong, W. 2022. Dynamics between regional sex ratios at birth and sex ratios at prime marriageable ages in China. *Population and Development Review* 48 (2):545–78. doi: 10.1111/padr.12476.

Yin, C., F. Meng, X. Yang, F. Yang, P. Fu, G. Yao, and R. Chen. 2022. Spatio-temporal evolution of urban built-up areas and analysis of driving factors—A comparison of typical cities in North and South China. *Land Use Policy* 117:106114. doi: 10.1016/j.landusepol.2022.106114.

TERENCE DAY is a Professor in the Department of Geography, Earth and Environmental Science at Okanagan College, Kelowna, BC V1Y 4X8, Canada. E-mail: tday@okanagan.bc.ca. He is a physical geographer with research interests in undergraduate teaching and learning, and the scope and nature of physical geography.