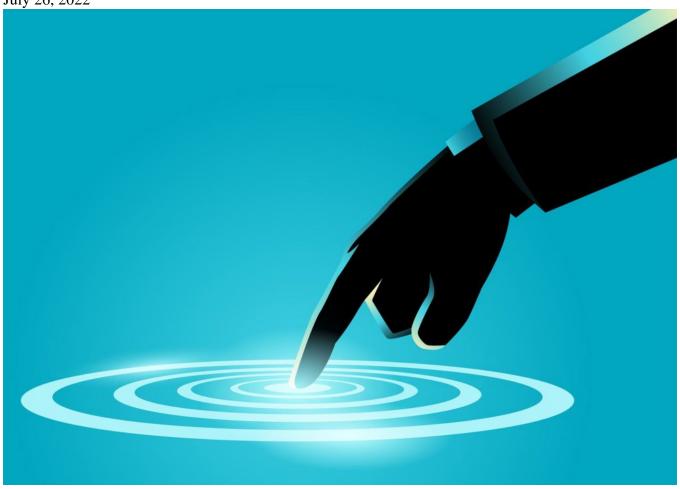# SAGE 10-Year Impact Awards Honor Two APS Articles

July 26, 2022

Two 2011 APS journal articles exploring the rise of Amazon Mechanical Turk (MTurk) and the risk of accepting false-positive findings have received SAGE Publishing's third annual 10-Year Impact Awards. With more than 11,600 citations between them, the two APS articles, along with a third published in the *Scandinavian Journal of Public Health*, are the most cited of all the studies SAGE published in 2011.

"[Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?](#)", by Michael Buhrmester, Tracy Kwang, and Samuel D. Gosling, appeared in *Perspectives on Psychological Science* and has more than 7,500 citations. "[False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant,](#)" by Joseph P. Simmons, Leif D. Nelson, and Uri Simonsohn, appeared in *Psychological Science* and has more than 4,100 citations.

"One size does not fit all in measuring research excellence, and relying solely on indicators that track citations two years later fails to capture the value that research can continue to bring—both inside and outside of academia—for far longer," said Ziyad Marar, SAGE's president of Global Publishing, in a [press release](#). "As a company committed to enabling the development and lasting impact of knowledge

for public good, we hope these awards will encourage a more nuanced array of metrics that capture the complex, subtle, diffuse, many-layered and—often—long-term nature of research impact."

[Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?](#)

MTurk), conceived as a crowdsourcing website for requesters to hire remotely located "crowdworkers" to perform discrete tasks, was launched publicly on November 2, 2005. Soon, psychological scientists started using the platform to collect data. However, concerns about data integrity and reliability emerged. In their 2011 article, Buhrmester, Kwang, and Gosling provided one of the first assessments of MTurk along with a description of its potential contributions to psychology and other social sciences.

The researchers analyzed the samples available on MTurk and how their performance reflected compensation. Buhrmester and colleagues concluded that MTurk participants were slightly more demographically diverse than standard internet samples and significantly more diverse than the typical participants in psychology experiments—American college students. Compensation rate and task length appeared to influence participation, the researchers reported, but participants could still be recruited rapidly and inexpensively. And when compensation rates were adequate for the task, the quality of the data was comparable to that of data collected in the lab or via traditional methods, and the reliability of the data was indistinct from that of traditionally collected data. The overall assessment was that MTurk could be used to quickly and inexpensively collect good-quality data.

"We anticipate that MTurk will soon become a major tool for research in psychology and elsewhere in the social sciences," Buhrmester and colleagues predicted in 2011. They were right.

Related: [Under the Hood of Mechanical Turk](#)

[False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant](#)

Accepting false-positive findings (i.e., obtaining a statistically significant p value, and thus rejecting a null hypothesis that is actually true) poses many risks to the healthy development of science. False positives are persistent once published, they waste resources by inspiring investment in fruitless research programs, and they can lead to ineffective policy changes, wrote Simmons, Nelson, and Simonsohn in their 2011 article. Moreover, "a field known for publishing false positives risks losing its credibility."

In their highly cited article, Simmons and colleagues showed that, in 2011, existing standards for disclosing details of data collection and analyses made false positives very likely and made it "unacceptably easy to publish 'statistically significant' evidence consistent with any hypothesis." The potential for researchers to make decisions regarding data collection as it happened and analysis and to explore different hypotheses after the fact—the *researcher degrees of freedom*—contributed to this inflation of false positives, they added.

Simmons and colleagues illustrated this problem with two experiments designed to demonstrate that certain songs can change listeners' age. Despite this "finding" being obviously wrong, their analyses yielded a false positive. They also provided a simulation illustrating how high the rate of false positives can be when researchers have different degrees of freedom.

Importantly, the authors provided a list of recommendations for researchers and reviewers to reduce false positives in psychological science. These recommendations included "authors must decide the rule for terminating data collection before data collection begins and report this rule in the article" and "authors must report all experimental conditions, including failed manipulations." In the years after the article's publication, open science practices have become increasingly common, with researchers becoming more transparent about their data collection and analysis. In 2014, APS started encouraging researchers to preregister their hypotheses and methods so that their degrees of freedom could be reduced and the rate of false positives in psychological science curtailed.

Related:
[Easier Done Than Said: Lessons From 6 Years of Preregistration](#)
[Seven Selfish Reasons for Preregistration](#)
[Preregistration Becoming the Norm in Psychological Science](#)

Feedback on this article? Email apsobserver@psychologicalscience.org or login to comment.

Works cited

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3–5. https://doi.org/10.1177/1745691610393980

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632