Retraction of a journal article doesn't make its findings false

October 1 2018, by Stephen S Holden

Wansink's research showed plate size matters when it comes to how much we eat. Credit: rawpixel/Unsplash, CC BY

The American Medical Association recently retracted six papers co-authored by food consumption and psychology researcher, Brian

Wansink, in three of its journals. These studies include two showing that large bowl sizes encourage us to eat more, and that shopping when hungry leads us to buy more calorie-dense foods.

A prolific academic researcher, Wansink has provided many thought-provoking ideas about the psychology of food consumption through more than 500 publications which have been collectively cited more than 25,000 times.

His research has shown that people will eat a lot more from a bottomless soup bowl; they will eat more from larger portions, even if it is stale popcorn or food served in a dark restaurant; and they will eat less if a portion is made to appear larger using visual illusions.

Retractions are a permanent means by which journals endeavour to preserve the integrity of scientific literature. They are typically issued for some form of misconduct, but it does not necessarily mean the results are false.

Are retracted studies false?

A number of challenges have been made against more than 50 of Wansink's publications. At present, 15 corrections have been published and 13 retractions have been made.

The retractions follow a range of allegations of misconduct including autoplagiarism (copying your own work), data mismanagement and data manipulation. But none of this means Wansink's results are entirely discredited.

The American Medical Association made its retractions based on Cornell University (Wansink's employer) being unable to provide an independent evaluation in response to an Expression of Concern

regarding Wansink's studies issued in May.

The absence of evidence does not prove his results are false.

Science relies far more on whether results are repeatable than retractions. And many of Wansink's results – including some which have been retracted – have been replicated.

Two of the most recently retracted studies showing that adults and children eat more from larger bowls form a part of a larger literature and have been cited nearly 300 times and 40 times respectively.

Multiple reviews of the scientific literature reveal that others have replicated the findings of Wansink and colleagues on how the plate or bowl size affects consumption.

In a meta-analysis I authored with others, the combined studies in this area show that doubling the plate size increases consumption by 40% on average. Though this is only the case if people are serving food onto the plate themselves. (Disclosure: this meta-analysis was published in a journal issue for which Wansink was one of the editors).

Replication is more important than retraction

The problem of reproducing findings in science is a much bigger issue than retractions. Retractions attract attention, but are relatively minor; replication does not attract attention, and is critically important.

The replication crisis facing social sciences, health and medicine suggests that 50% or more of published findings may not be repeatable.

In social science, a team replicated 100 studies published in three high-ranking journals. The results showed only 36% of the replications found

statistically significant results, and the average size of the observed effects was half of that seen in the original studies.

Wansink has published more than 500 articles. If 250 of them prove to be false in the sense that the results cannot be replicated, then he is on par with social and medical science in general.

The retraction of thirteen of Wansink's articles—some of which have been replicated by others—is a blip receiving much more attention than it deserves.

The high rate of replication failure arises, in part, from the arcane statistical approach used for analysing research data. In essence, researchers seek statistically significant findings. Statistical significance is typically defined as when the probability (p-value) of the observed data *assuming there was no effect* is less than 5%.

Journals and academics wish to publish novel, statistically significant results. They tend to ignore studies with null results, putting them in a file-drawer.

Replications that are successful add nothing new, and replications that fail (not statistically significant) are uninteresting to publishers albeit critically important to science.

A related problem is that academics may dredge through data and cherry pick statistically significant results, a practice called p-hacking.

The misconduct of journals and academics through their obsessive focus on statistically significant findings is widespread. If Wansink differs from others, it is in his disarming honesty admitting to data dredging in a 2016 blog post which attracted intensive scrutiny from his peers.

Science makes mistakes and missteps. The advances are achieved through new ideas and repeated testing.

Retractions may be important signals of reduced confidence in a finding, but they do not prove a finding false. This requires replication.

Science doesn't provide certainty. Claims of absolute certainty made by authoritative figures are probably false.

As Tim van der Zee, one of Wansink's lead detractors states on his website "I am wrong most of the time." The challenge for scientists is to believe this.

Provided by The Conversation