

Replication crisis

The **replication crisis** (also called the **replicability crisis** and the **reproducibility crisis**) is an ongoing methodological crisis in which the results of many scientific studies are difficult or impossible to reproduce. Because the reproducibility of empirical results is an essential part of the scientific method,^[2] such failures undermine the credibility of theories building on them and potentially call into question substantial parts of scientific knowledge.

The replication crisis is frequently discussed in relation to psychology and medicine, where considerable efforts have been undertaken to reinvestigate classic results, to determine whether they are reliable, and if they turn out not to be, the reasons for the failure.^{[3][4]} Data strongly indicates that other natural and social sciences are affected as well.^[5]

The phrase *replication crisis* was coined in the early 2010s^[6] as part of a growing awareness of the problem. Considerations of causes and remedies have given rise to a new scientific discipline, metascience,^[7] which uses methods of empirical research to examine empirical research practice.

Considerations about reproducibility fall into two categories. *Reproducibility* in the narrow sense refers to re-examining and validating the analysis of a given set of data. *Replication* refers to repeating the experiment or study to obtain new, independent data with the goal of reaching the same or similar conclusions.

Background

Replication has been called "the cornerstone of science".^{[8][9]} Environmental health scientist Stefan Schmidt began a 2009 review with this description of replication:

Replication is one of the central issues in any empirical science. To confirm results or hypotheses by a repetition procedure is at the basis of any scientific conception. A replication experiment to demonstrate that the same findings can be obtained in any other place by any other researcher is conceived as an operationalization of objectivity. It is the proof that the experiment reflects knowledge that can be separated from the specific circumstances (such as time, place, or persons) under which it was gained.^[10]

But there is limited consensus on how to define *replication* and potentially related concepts.^[11]^{[12][10]} A number of types of replication have been identified:



Ioannidis (2005): "Why Most Published Research Findings Are False".^[1]

1. *Direct or exact replication*, where an experimental procedure is repeated as closely as possible.^{[10][13]}
2. *Systematic replication*, where an experimental procedure is largely repeated, with some intentional changes.^[13]
3. *Conceptual replication*, where a finding or hypothesis is tested using a different procedure.^{[10][13]} Conceptual replication allows testing for generalizability and veracity of a result or hypothesis.^[13]

Reproducibility can also be distinguished from *replication*, as referring to reproducing the same results using the same data set. Reproducibility of this type is why many researchers make their data available to others for testing.^[14]

The replication crisis does not necessarily mean these fields are unscientific.^{[15][16][17]} Rather, this process is part of the scientific process in which old ideas or those that cannot withstand careful scrutiny are pruned,^{[18][19]} although this pruning process is not always effective.^{[20][21]}

A hypothesis is generally considered to be supported when the results match the predicted pattern and that pattern of results is found to be statistically significant. Results are considered significant whenever the relative frequency of the observed pattern falls below an arbitrarily chosen value (i.e. the significance level) when assuming the null hypothesis is true. This generally answers the question of how unlikely results would be if no difference existed at the level of the statistical population. If the probability associated with the test statistic exceeds the chosen critical value, the results are considered statistically significant.^[22] The corresponding probability of exceeding the critical value is depicted as $p < 0.05$, where p (typically referred to as the "p-value") is the probability level. This should result in 5% of hypotheses that are supported being false positives (an incorrect hypothesis being erroneously found correct), assuming the studies meet all of the statistical assumptions. Some fields use smaller p-values, such as $p < 0.01$ (1% chance of a false positive) or $p < 0.001$ (0.1% chance of a false positive). But a smaller chance of a false positive often requires greater sample sizes or a greater chance of a false negative (a correct hypothesis being erroneously found incorrect). Although p -value testing is the most commonly used method, it is not the only method.

History

The beginning of the replication crisis can be traced to a number of events in the early 2010s. Philosopher of science and social epistemologist Felipe Romero identified four events that can be considered precursors to the ongoing crisis:^[23]

- **Controversies around social priming research:** In the early 2010s, the well-known "elderly-walking" study^[24] by social psychologist John Bargh and colleagues failed to replicate in two direct replications.^[25] This experiment was part of a series of three studies that had been widely cited throughout the years, was regularly taught in university courses, and had inspired a large number of conceptual replications. Failures to replicate the study led to much controversy and a heated debate involving the original authors.^[26] Notably, many of the conceptual replications of the original studies also failed to replicate in subsequent direct replications.^{[27][28][29][30]}
- **Controversies around experiments on extrasensory perception:** Social psychologist Daryl Bem conducted a series of experiments supposedly providing evidence for the controversial phenomenon of extrasensory perception.^[31] Bem was highly criticized for his study's methodology and upon reanalysis of the data, no evidence was found for the

existence of extrasensory perception.^[32] The experiment also failed to replicate in subsequent direct replications.^[33] According to Romero, what the community found particularly upsetting was that many of the flawed procedures and statistical tools used in Bem's studies were part of common research practice in psychology.

- **Amgen and Bayer reports on lack of replicability in biomedical research:** Scientists from biotech companies [Amgen](#) and [Bayer Healthcare](#) reported alarmingly low replication rates (11–20%) of landmark findings in preclinical oncological research.^{[34][35]}
- **Publication of studies on p-hacking and questionable research practices:** Since the late 2000s, a number of studies in [metascience](#) showed how commonly adopted practices in many scientific fields, such as exploiting the flexibility of the process of data collection and reporting, could greatly increase the probability of false positive results.^{[36][37][38]} These studies suggested how a significant proportion of published literature in several scientific fields could be nonreplicable research.

This series of events generated a great deal of skepticism about the validity of existing research in light of widespread methodological flaws and failures to replicate findings. This led prominent scholars to declare a "crisis of confidence" in psychology and other fields,^[39] and the ensuing situation came to be known as the "replication crisis".

Although the beginning of the replication crisis can be traced to the early 2010s, some authors point out that concerns about replicability and research practices in the social sciences had been expressed much earlier. Romero notes that authors voiced concerns about the lack of direct replications in psychological research in the late 1960s and early 1970s.^{[40][41]} He also writes that certain studies in the 1990s were already reporting that journal editors and reviewers are generally biased against publishing replication studies.^{[42][43]}

In the social sciences, the blog [Data Colada](#) (whose three authors coined the term "p-hacking" in a 2014 paper) has been credited with contributing to the start of the replication crisis.^{[44][45][46]}

University of Virginia professor and cognitive psychologist [Barbara A. Spellman](#) has written that many criticisms of research practices and concerns about replicability of research are not new.^[47] She reports that between the late 1950s and the 1990s, scholars were already expressing concerns about a possible crisis of replication,^[48] a suspiciously high rate of positive findings,^[49] questionable research practices (QRPs),^[50] the effects of publication bias,^[51] issues with statistical power,^{[52][53]} and bad standards of reporting.^[48]

Spellman also identifies reasons that the reiteration of these criticisms and concerns in recent years led to a full-blown crisis and challenges to the status quo. First, technological improvements facilitated conducting and disseminating replication studies, and analyzing large swaths of literature for systemic problems. Second, the research community's increasing size and diversity made the work of established members more easily scrutinized by other community members unfamiliar with them. According to Spellman, these factors, coupled with increasingly limited resources and misaligned incentives for doing scientific work, led to a crisis in psychology and other fields.^[47]

Prevalence

In psychology

Several factors have combined to put psychology at the center of the conversation.^{[54][55]} Some

areas of psychology once considered solid, such as social priming, have come under increased scrutiny due to failed replications.^[56] Much of the focus has been on the area of social psychology,^[57] although other areas of psychology such as clinical psychology,^{[58][59][60]} developmental psychology,^{[61][62][63]} and educational research have also been implicated.^{[64][65][66][67][68]}

In August 2015, the first open empirical study of reproducibility in psychology was published, called The Reproducibility Project: Psychology. Coordinated by psychologist Brian Nosek, researchers redid 100 studies in psychological science from three high-ranking psychology journals (*Journal of Personality and Social Psychology*, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, and *Psychological Science*). 97 of the original studies had significant effects, but of those 97, only 36% of the replications yielded significant findings (p value below 0.05).^[41] The mean effect size in the replications was approximately half the magnitude of the effects reported in the original studies. The same paper examined the reproducibility rates and effect sizes by journal and discipline. Study replication rates were 23% for the *Journal of Personality and Social Psychology*, 48% for *Journal of Experimental Psychology: Learning, Memory, and Cognition*, and 38% for *Psychological Science*. Studies in the field of cognitive psychology had a higher replication rate (50%) than studies in the field of social psychology (25%).^[69]

A study published in 2018 in Nature Human Behaviour replicated 21 social and behavioral science papers from *Nature* and *Science*, finding that only about 62% could successfully reproduce original results.^{[70][71]}

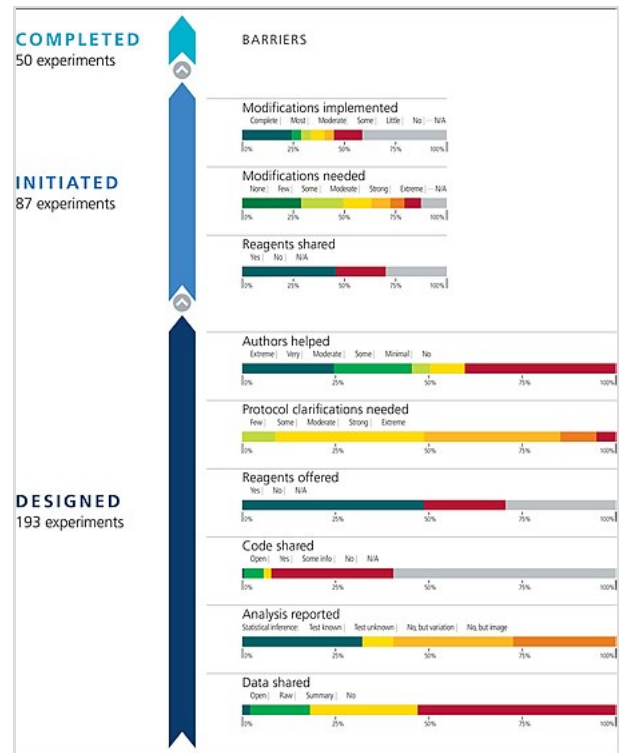
Similarly, in a study conducted under the auspices of the Center for Open Science, a team of 186 researchers from 60 different laboratories (representing 36 different nationalities from six different continents) conducted replications of 28 classic and contemporary findings in psychology.^{[72][73]} The study's focus was not only whether the original papers' findings replicated but also the extent to which findings varied as a function of variations in samples and contexts. Overall, 50% of the 28 findings failed to replicate despite massive sample sizes. But if a finding replicated, then it replicated in most samples. If a finding was not replicated, then it failed to replicate with little variation across samples and contexts. This evidence is inconsistent with a proposed explanation that failures to replicate in psychology are likely due to changes in the sample between the original and replication study.^[73]

Results of a 2022 study suggest that many earlier brain–phenotype studies ("brain-wide association studies" (BWAS)) produced invalid conclusions as the replication of such studies requires samples from thousands of individuals due to small effect sizes.^{[74][75]}

In medicine

Of 49 medical studies from 1990 to 2003 with more than 1000 citations, 92% found that the studied therapies were effective. Of these studies, 16% were contradicted by subsequent studies, 16% had found stronger effects than did subsequent studies, 44% were replicated, and 24% remained largely unchallenged.^[76] A 2011 analysis by researchers with pharmaceutical company Bayer found that, at most, a quarter of Bayer's in-house findings replicated the original results.^[77] But the analysis of Bayer's results found that the results that did replicate could often be successfully used for clinical applications.^[78]

In a 2012 paper, C. Glenn Begley, a biotech consultant working at Amgen, and Lee Ellis, a medical researcher at the University of Texas, found that only 11% of 53 pre-clinical cancer studies had replications that could confirm conclusions from the original studies.^[79] In late 2021, The Reproducibility Project: Cancer Biology examined 53 top papers about cancer published between 2010 and 2012 and showed that among studies that provided sufficient information to be redone, the effect sizes were 85% smaller on average than the original findings.^{[80][81]} A survey of cancer researchers found that half of them had been unable to reproduce a published result.^[82] Another report estimated that almost half of randomized controlled trials contained flawed data (based on the analysis of anonymized individual participant data (IPD) from more than 150 trials).^[83]



Results from The Reproducibility Project: Cancer Biology suggest most studies of the cancer research sector may not be replicable.

In other disciplines

In economics

Economics has lagged behind other social sciences and psychology in its attempts to assess replication rates and increase the number of studies that attempt replication.^[12] A 2016 study in the journal *Science* replicated 18 experimental studies published in two leading economics journals, *The American Economic Review* and the *Quarterly Journal of Economics*, between 2011 and 2014. It found that about 39% failed to reproduce the original results.^{[84][85][86]} About 20% of studies published in *The American Economic Review* are contradicted by other studies despite relying on the same or similar data sets.^[87] A study of empirical findings in the *Strategic Management Journal* found that about 30% of 27 retested articles showed statistically insignificant results for previously significant findings, whereas about 4% showed statistically significant results for previously insignificant findings.^[88]

In water resource management

A 2019 study in *Scientific Data* estimated with 95% confidence that of 1,989 articles on water resources and management published in 2017, study results might be reproduced for only 0.6% to 6.8%, even if each of these articles were to provide sufficient information that allowed for replication.^[89]

Across fields

A 2016 survey by *Nature* on 1,576 researchers who took a brief online questionnaire on reproducibility found that more than 70% of researchers have tried and failed to reproduce another scientist's experiment results (including 87% of chemists, 77% of biologists, 69% of

physicists and engineers, 67% of medical researchers, 64% of earth and environmental scientists, and 62% of all others), and more than half have failed to reproduce their own experiments. But fewer than 20% had been contacted by another researcher unable to reproduce their work. The survey found that fewer than 31% of researchers believe that failure to reproduce results means that the original result is probably wrong, although 52% agree that a significant replication crisis exists. Most researchers said they still trust the published literature.^{[5][90]} In 2010, Fanelli (2010)^[91] found that 91.5% of psychiatry/psychology studies confirmed the effects they were looking for, and concluded that the odds of this happening (a positive result) was around five times higher than in fields such as astronomy or geosciences. Fanelli argued that this is because researchers in "softer" sciences have fewer constraints to their conscious and unconscious biases.

Early analysis of result-blind peer review, which is less affected by publication bias, has estimated that 61% of result-blind studies in biomedicine and psychology have led to null results, in contrast to an estimated 5% to 20% in earlier research.^[92]

Causes

The replication crisis may be triggered by the "generation of new data and scientific publications at an unprecedented rate" that leads to the "desperation to publish or perish" and a failure to adhere to good scientific practice.^[93]

Historical and sociological roots

Predictions of an impending crisis in the quality control mechanism of science can be traced back several decades. Derek de Solla Price—considered the father of scientometrics, the quantitative study of science—predicted in 1963 that science could reach "senility" as a result of its own exponential growth.^[94] Some present-day literature seems to vindicate this "overflow" prophecy, lamenting the decay in both attention and quality.^{[95][96]}

Historian Philip Mirowski argues that the decline of scientific quality can be connected to its commodification, especially spurred by the profit-driven decision of major corporations to outsource their research to universities and contract research organizations.^[97]

Social systems theory, as expounded in the work of German sociologist Niklas Luhmann, inspires a similar diagnosis. This theory holds that each system, such as economy, science, religion or media, communicates using its own code: *true* and *false* for science, *profit* and *loss* for the economy, *news* and *no-news* for the media, and so on.^{[98][99]} According to some sociologists, science's mediatization,^[100] its commodification^[97] and its politicization,^{[100][101]} as a result of the structural coupling among systems, have led to a confusion of the original system codes.

Problems with the publication system in science

Publication bias

A major cause of low reproducibility is the publication bias stemming from the fact that statistically non-significant results and seemingly unoriginal replications are rarely published.

Only a very small proportion of academic journals in psychology and neurosciences explicitly welcomed submissions of replication studies in their aim and scope or instructions to authors.^{[102][103]} This does not encourage reporting on, or even attempts to perform, replication studies. Among 1,576 researchers *Nature* surveyed in 2016, only a minority had ever attempted to publish a replication, and several respondents who had published failed replications noted that editors and reviewers demanded that they play down comparisons with the original studies.^{[5][90]} An analysis of 4,270 empirical studies in 18 business journals from 1970 to 1991 reported that less than 10% of accounting, economics, and finance articles and 5% of management and marketing articles were replication studies.^{[84][104]} Publication bias is augmented by the pressure to publish and the author's own confirmation bias,^[a] and is an inherent hazard in the field, requiring a certain degree of skepticism on the part of readers.^[38]

Publication bias leads to what psychologist Robert Rosenthal calls the "file drawer effect". The file drawer effect is the idea that as a consequence of the publication bias, a significant number of negative results^[b] are not published. According to philosopher of science Felipe Romero, this tends to produce "misleading literature and biased meta-analytic studies",^[23] and when publication bias is considered along with the fact that a majority of tested hypotheses might be false *a priori*, it is plausible that a considerable proportion of research findings might be false positives, as shown by metascientist John Ioannidis.^[1] In turn, a high proportion of false positives in the published literature can explain why many findings are nonreproducible.^[23]

"Publish or perish" culture

The consequences for replicability of the publication bias are exacerbated by academia's "publish or perish" culture. As explained by metascientist Daniele Fanelli, "publish or perish" culture is a sociological aspect of academia whereby scientists work in an environment with very high pressure to have their work published in recognized journals. This is the consequence of the academic work environment being hypercompetitive and of bibliometric parameters (e.g., number of publications) being increasingly used to evaluate scientific careers.^[106] According to Fanelli, this pushes scientists to employ a number of strategies aimed at making results "publishable". In the context of publication bias, this can mean adopting behaviors aimed at making results positive or statistically significant, often at the expense of their validity (see QRPs, section 4.3).^[106]

According to Center for Open Science founder Brian Nosek and his colleagues, "publish or perish" culture created a situation whereby the goals and values of single scientists (e.g., publishability) are not aligned with the general goals of science (e.g., pursuing scientific truth). This is detrimental to the validity of published findings.^[107]

Philosopher Brian D. Earp and psychologist Jim A. C. Everett argue that, although replication is in the best interests of academics and researchers as a group, features of academic psychological culture discourage replication by individual researchers. They argue that performing replications can be time-consuming, and take away resources from projects that reflect the researcher's original thinking. They are harder to publish, largely because they are unoriginal, and even when they can be published they are unlikely to be viewed as major contributions to the field. Replications "bring less recognition and reward, including grant money, to their authors".^[108]

In his 1971 book *Scientific Knowledge and Its Social Problems*, philosopher and historian of

science Jerome R. Ravetz predicted that science—in its progression from "little" science composed of isolated communities of researchers to "big" science or "techno-science"—would suffer major problems in its internal system of quality control. He recognized that the incentive structure for modern scientists could become dysfunctional, creating perverse incentives to publish any findings, however dubious. According to Ravetz, quality in science is maintained only when there is a community of scholars, linked by a set of shared norms and standards, who are willing and able to hold each other accountable.

Standards of reporting

Certain publishing practices also make it difficult to conduct replications and to monitor the severity of the reproducibility crisis, for articles often come with insufficient descriptions for other scholars to reproduce the study. The Reproducibility Project: Cancer Biology showed that of 193 experiments from 53 top papers about cancer published between 2010 and 2012, only 50 experiments from 23 papers have authors who provided enough information for researchers to redo the studies, sometimes with modifications. None of the 193 papers examined had its experimental protocols fully described and replicating 70% of experiments required asking for key reagents.^{[80][81]} The aforementioned study of empirical findings in the *Strategic Management Journal* found that 70% of 88 articles could not be replicated due to a lack of sufficient information for data or procedures.^{[84][88]} In water resources and management, most of 1,987 articles published in 2017 were not replicable because of a lack of available information shared online.^[89]

Questionable research practices and fraud

Questionable research practices (QRPs) are intentional behaviors that capitalize on the gray area of acceptable scientific behavior or exploit the researcher degrees of freedom (researcher DF), which can contribute to the irreproducibility of results by increasing the probability of false positive results.^{[109][110][38]} Researcher DF are seen in hypothesis formulation, design of experiments, data collection and analysis, and reporting of research.^[110] Some examples of QRPs are data dredging,^{[110][111][37][c]} selective reporting,^{[109][110][111][37][d]} and HARKing (hypothesising after results are known).^{[110][111][37][e]} In medicine, irreproducible studies have six features in common. These include investigators not being blinded to the experimental versus the control arms, a failure to repeat experiments, a lack of positive and negative controls, failing to report all the data, inappropriate use of statistical tests, and use of reagents that were not appropriately validated.^[113]

QRPs do not include more explicit violations of scientific integrity, such as data falsification.^{[109][110]} Fraudulent research does occur, as in the case of scientific fraud by social psychologist Diederik Stapel,^{[114][13]} cognitive psychologist Marc Hauser and social psychologist Lawrence Sanna,^[13] but it appears to be uncommon.^[13]

Prevalence

According to IU professor Ernest O'Boyle and psychologist Martin Götz, around 50% of researchers surveyed across various studies admitted engaging in HARKing.^[115] In a survey of 2,000 psychologists by behavioral scientist Leslie K. John and colleagues, around 94% of psychologists admitted having employed at least one QRP. More specifically, 63% admitted failing to report all of a study's dependent measures, 28% to report all of a study's conditions,

and 46% to selectively reporting studies that produced the desired pattern of results. In addition, 56% admitted having collected more data after having inspected already collected data, and 16% to having stopped data collection because the desired result was already visible.^[37] According to biotechnology researcher J. Leslie Glick's estimate in 1992, 10% to 20% of research and development studies involved either QRPs or outright fraud.^[116] The methodology used to estimate QRPs has been contested, and more recent studies suggested lower prevalence rates on average.^[117]

A 2009 meta-analysis found that 2% of scientists across fields admitted falsifying studies at least once and 14% admitted knowing someone who did. Such misconduct was, according to one study, reported more frequently by medical researchers than by others.^[118]

Statistical issues

Low statistical power

According to Deakin University professor Tom Stanley and colleagues, one plausible reason studies fail to replicate is low statistical power. This happens for three reasons. First, a replication study with low power is unlikely to succeed since, by definition, it has a low probability to detect a true effect. Second, if the original study has low power, it will yield biased effect size estimates. When conducting a priori power analysis for the replication study, this will result in underestimation of the required sample size. Third, if the original study has low power, the post-study odds of a statistically significant finding reflecting a true effect are quite low. It is therefore likely that a replication attempt of the original study would fail.^[14]

Stanley and colleagues estimated the average statistical power of psychological literature by analyzing data from 200 meta-analyses. They found that on average, psychology studies have between 33.1% and 36.4% statistical power. These values are quite low compared to the 80% considered adequate statistical power for an experiment. Across the 200 meta-analyses, the median of studies with adequate statistical power was between 7.7% and 9.1%.^[14]

In a study published in *Nature*, psychologist Katherine Button and colleagues conducted a similar study with 49 meta-analyses in neuroscience, estimating a median statistical power of 21%.^[119] Meta-scientist John Ioannidis and colleagues computed an estimate of average power for empirical economic research, finding a median power of 18% based on literature drawing upon 6.700 studies.^[120] In light of these results, it is plausible that a major reason for widespread failures to replicate in several scientific fields might be very low statistical power on average.

Statistical heterogeneity

As also reported by Stanley and colleagues, a further reason studies might fail to replicate is high heterogeneity of the to-be-replicated effects. In meta-analysis, "heterogeneity" refers to the variance in research findings that results from there being no single true effect size. Instead, findings in such cases are better seen as a distribution of true effects.^[14] Statistical heterogeneity is calculated using the I-squared statistic,^[121] defined as "the proportion (or percentage) of observed variation among reported effect sizes that cannot be explained by the calculated standard errors associated with these reported effect sizes".^[14] This variation can be

due to differences in experimental methods, populations, cohorts, and statistical methods between replication studies. Heterogeneity poses a challenge to studies attempting to replicate previously found effect sizes. When heterogeneity is high, subsequent replications have a high probability of finding an effect size radically different than that of the original study.^[f]

Importantly, significant levels of heterogeneity are also found in direct/exact replications of a study. Stanley and colleagues discuss this while reporting a study by quantitative behavioral scientist Richard Klein and colleagues, where the authors attempted to replicate 15 psychological effects across 36 different sites in Europe and the U.S. In the study, Klein and colleagues found significant amounts of heterogeneity in 8 out of 16 effects (I-squared = 23% to 91%). Importantly, while the replication sites intentionally differed on a variety of characteristics, such differences could account for very little heterogeneity. According to Stanley and colleagues, this suggested that heterogeneity could have been a genuine characteristic of the phenomena being investigated. For instance, phenomena might be influenced by so-called "hidden moderators" – relevant factors that were previously not understood to be important in the production of a certain effect.

In their analysis of 200 meta-analyses of psychological effects, Stanley and colleagues found a median percent of heterogeneity of I-squared = 74%. According to the authors, this level of heterogeneity can be considered "huge". It is three times larger than the random sampling variance of effect sizes measured in their study. If considered along sampling error, heterogeneity yields a standard deviation from one study to the next even larger than the median effect size of the 200 meta-analyses they investigated.^[g] The authors conclude that if replication is defined by a subsequent study finding a sufficiently similar effect size to the original, replication success is not likely even if replications have very large sample sizes. Importantly, this occurs even if replications are direct or exact since heterogeneity nonetheless remains relatively high in these cases.

Others

Within economics, the replication crisis may be also exacerbated because econometric results are fragile.^[122] using different but plausible estimation procedures or data preprocessing techniques can lead to conflicting results.^{[123][124][125]}

Context sensitivity

New York University professor Jay Van Bavel and colleagues argue that a further reason findings are difficult to replicate is the sensitivity to context of certain psychological effects. On this view, failures to replicate might be explained by contextual differences between the original experiment and the replication, often called "hidden moderators".^[126] Van Bavel and colleagues tested the influence of context sensitivity by reanalyzing the data of the widely cited Reproducibility Project carried out by the Open Science Collaboration.^[11] They re-coded effects according to their sensitivity to contextual factors and then tested the relationship between context sensitivity and replication success in various regression models.

Context sensitivity was found to negatively correlate with replication success, such that higher ratings of context sensitivity were associated with lower probabilities of replicating an effect.^[h] Importantly, context sensitivity significantly correlated with replication success even when adjusting for other factors considered important for reproducing results (e.g., effect size and

sample size of original, statistical power of the replication, methodological similarity between original and replication).^[1] In light of the results, the authors concluded that attempting a replication in a different time, place or with a different sample can significantly alter an experiment's results. Context sensitivity thus may be a reason certain effects fail to replicate in psychology.^[126]

Base rate of hypothesis accuracy

According to philosopher Alexander Bird, a possible reason for the low rates of replicability in certain scientific fields is that a majority of tested hypotheses are false *a priori*.^[127] On this view, low rates of replicability could be consistent with quality science. Relatedly, the expectation that most findings should replicate would be misguided and, according to Bird, a form of base rate fallacy. Bird's argument works as follows. Assuming an ideal situation of a test of significance, whereby the probability of incorrectly rejecting the null hypothesis is 5% (i.e. Type I error) and the probability of correctly rejecting the null hypothesis is 80% (i.e. Power), in a context where a high proportion of tested hypotheses are false, it is conceivable that the number of false positives would be high compared to those of true positives.^[127] For example, in a situation where only 10% of tested hypotheses are actually true, one can calculate that as much as 36% of results will be false positives.^[1]

The claim that the falsity of most tested hypotheses can explain low rates of replicability is even more relevant when considering that the average power for statistical tests in certain fields might be much lower than 80%. For example, the proportion of false positives increases to a value between 55.2% and 57.6% when calculated with the estimates of an average power between 34.1% and 36.4% for psychology studies, as provided by Stanley and colleagues in their analysis of 200 meta-analyses in the field.^[14] A high proportion of false positives would then result in many research findings being non-replicable.

Bird notes that the claim that a majority of tested hypotheses are false *a priori* in certain scientific fields might be plausible given factors such as the complexity of the phenomena under investigation, the fact that theories are seldom undisputed, the "inferential distance" between theories and hypotheses, and the ease with which hypotheses can be generated. In this respect, the fields Bird takes as examples are clinical medicine, genetic and molecular epidemiology, and social psychology. This situation is radically different in fields where theories have outstanding empirical basis and hypotheses can be easily derived from theories (e.g., experimental physics).^[127]

Consequences

When effects are wrongly stated as relevant in the literature, failure to detect this by replication will lead to the canonization of such false facts.^[128]

A 2021 study found that papers in leading general interest, psychology and economics journals with findings that could not be replicated tend to be cited more over time than reproducible research papers, likely because these results are surprising or interesting. The trend is not affected by publication of failed reproductions, after which only 12% of papers that cite the original research will mention the failed replication.^{[129][130]} Further, experts are able to predict which studies will be replicable, leading the authors of the 2021 study, Marta Serra-Garcia and Uri Gneezy, to conclude that experts apply lower standards to interesting results when deciding

whether to publish them.^[130]

Public awareness and perceptions

Concerns have been expressed within the scientific community that the general public may consider science less credible due to failed replications.^[131] Research supporting this concern is sparse, but a nationally representative survey in Germany showed that more than 75% of Germans have not heard of replication failures in science.^[132] The study also found that most Germans have positive perceptions of replication efforts: only 18% think that non-replicability shows that science cannot be trusted, while 65% think that replication research shows that science applies quality control, and 80% agree that errors and corrections are part of science.^[132]

Response in academia

With the replication crisis of psychology earning attention, Princeton University psychologist Susan Fiske drew controversy for speaking against critics of psychology for what she called bullying and undermining the science.^{[133][134][135][136]} She called these unidentified "adversaries" names such as "methodological terrorist" and "self-appointed data police", saying that criticism of psychology should be expressed only in private or by contacting the journals.^[133] Columbia University statistician and political scientist Andrew Gelman responded to Fiske, saying that she had found herself willing to tolerate the "dead paradigm" of faulty statistics and had refused to retract publications even when errors were pointed out.^[133] He added that her tenure as editor had been abysmal and that a number of published papers she edited were found to be based on extremely weak statistics; one of Fiske's own published papers had a major statistical error and "impossible" conclusions.^[133]

Credibility revolution

Some researchers in psychology indicate that the replication crisis is a foundation for a "credibility revolution", where changes in standards by which psychological science are evaluated may include emphasizing transparency and openness, preregistering research projects, and replicating research with higher standards for evidence to improve the strength of scientific claims.^[137] Such changes may diminish the productivity of individual researchers, but this effect could be avoided by data sharing and greater collaboration.^[137] A credibility revolution could be good for the research environment.^[138]

Remedies

Focus on the replication crisis has led to renewed efforts in psychology to retest important findings.^{[38][139]} A 2013 special edition of the journal *Social Psychology* focused on replication studies.^[12]

Standardization as well as (requiring) transparency of the used statistical and experimental methods have been proposed.^[140] Careful documentation of the experimental set-up is considered crucial for replicability of experiments and various variables may not be documented and standardized such as animals' diets in animal studies.^[141]

A 2016 article by John Ioannidis elaborated on "Why Most Clinical Research Is Not Useful".^[142]

Ioannidis describes what he views as some of the problems and calls for reform, characterizing certain points for medical research to be useful again; one example he makes is the need for medicine to be patient-centered (e.g. in the form of the Patient-Centered Outcomes Research Institute) instead of the current practice to mainly take care of "the needs of physicians, investigators, or sponsors".

Reform in scientific publishing

Metascience

Metascience is the use of scientific methodology to study science itself. It seeks to increase the quality of scientific research while reducing waste. It is also known as "research on research" and "the science of science", as it uses research methods to study how research is done and where improvements can be made. Metascience is concerned with all fields of research and has been called "a bird's eye view of science."^[143] In Ioannidis's words, "Science is the best thing that has happened to human beings ... but we can do it better."^[144]

Meta-research continues to be conducted to identify the roots of the crisis and to address them. Methods of addressing the crisis include pre-registration of scientific studies and clinical trials as well as the founding of organizations such as CONSORT and the EQUATOR Network that issue guidelines for methodology and reporting. Efforts continue to reform the system of academic incentives, improve the peer review process, reduce the misuse of statistics, combat bias in scientific literature, and increase the overall quality and efficiency of the scientific process.

Presentation of methodology

Some authors have argued that the insufficient communication of experimental methods is a major contributor to the reproducibility crisis and that better reporting of experimental design and statistical analyses would improve the situation. These authors tend to plead for both a broad cultural change in the scientific community of how statistics are considered and a more coercive push from scientific journals and funding bodies.^[145] But concerns have been raised about the potential for standards for transparency and replication to be misapplied to qualitative as well as quantitative studies.^[146]

Business and management journals that have introduced editorial policies on data accessibility, replication, and transparency include the *Strategic Management Journal*, the *Journal of International Business Studies*, and the *Management and Organization Review*.^[84]

Result-blind peer review

In response to concerns in psychology about publication bias and data dredging, more than 140 psychology journals have adopted result-blind peer review. In this approach, studies are accepted not on the basis of their findings and after the studies are completed, but before they are conducted and on the basis of the methodological rigor of their experimental designs, and the theoretical justifications for their statistical analysis techniques before data collection or analysis is done.^[147] Early analysis of this procedure has estimated that 61% of result-blind studies have led to null results, in contrast to an estimated 5% to 20% in earlier research.^[92] In

addition, large-scale collaborations between researchers working in multiple labs in different countries that regularly make their data openly available for different researchers to assess have become much more common in psychology.^[148]

Pre-registration of studies

Scientific publishing has begun using pre-registration reports to address the replication crisis.^{[149][150]} The registered report format requires authors to submit a description of the study methods and analyses prior to data collection. Once the method and analysis plan is vetted through peer-review, publication of the findings is provisionally guaranteed, based on whether the authors follow the proposed protocol. One goal of registered reports is to circumvent the publication bias toward significant findings that can lead to implementation of questionable research practices. Another is to encourage publication of studies with rigorous methods.

The journal *Psychological Science* has encouraged the preregistration of studies and the reporting of effect sizes and confidence intervals.^[151] The editor in chief also noted that the editorial staff will be asking for replication of studies with surprising findings from examinations using small sample sizes before allowing the manuscripts to be published.

Metadata and digital tools for tracking replications

It has been suggested that "a simple way to check how often studies have been repeated, and whether or not the original findings are confirmed" is needed.^[129] Categorizations and ratings of reproducibility at the study or results level, as well as addition of links to and rating of third-party confirmations, could be conducted by the peer-reviewers, the scientific journal, or by readers in combination with novel digital platforms or tools.

Statistical reform

Requiring smaller p -values

Many publications require a p -value of $p < 0.05$ to claim statistical significance. The paper "Redefine statistical significance",^[152] signed by a large number of scientists and mathematicians, proposes that in "fields where the threshold for defining statistical significance for new discoveries is $p < 0.05$, we propose a change to $p < 0.005$. This simple step would immediately improve the reproducibility of scientific research in many fields." Their rationale is that "a leading cause of non-reproducibility (is that the) statistical standards of evidence for claiming new discoveries in many fields of science are simply too low. Associating 'statistically significant' findings with $p < 0.05$ results in a high rate of false positives even in the absence of other experimental, procedural and reporting problems."^[152]

This call was subsequently criticised by another large group, who argued that "redefining" the threshold would not fix current problems, would lead to some new ones, and that in the end, all thresholds needed to be justified case-by-case instead of following general conventions.^[153]

Addressing misinterpretation of p -values

Although statisticians are unanimous that the use of " $p < 0.05$ " as a standard for significance provides weaker evidence than is generally appreciated, there is a lack of unanimity about what

should be done about it. Some have advocated that Bayesian methods should replace p -values. This has not happened on a wide scale, partly because it is complicated and partly because many users distrust the specification of prior distributions in the absence of hard data. A simplified version of the Bayesian argument, based on testing a point null hypothesis was suggested by pharmacologist David Colquhoun.^{[154][155]} The logical problems of inductive inference were discussed in "The Problem with p -values" (2016).^[156]

The hazards of reliance on p -values arises partly because even an observation of $p = 0.001$ is not necessarily strong evidence against the null hypothesis.^[155] Despite the fact that the likelihood ratio in favor of the alternative hypothesis over the null is close to 100, if the hypothesis was implausible, with a prior probability of a real effect being 0.1, even the observation of $p = 0.001$ would have a false positive risk of 8 percent. It would still fail to reach the 5 percent level.

It was recommended that the terms "significant" and "non-significant" should not be used.^[155] p -values and confidence intervals should still be specified, but they should be accompanied by an indication of the false-positive risk. It was suggested that the best way to do this is to calculate the prior probability that would be necessary to believe in order to achieve a false positive risk of a certain level, such as 5%. The calculations can be done with various computer software.^{[155][157]} This reverse Bayesian approach, which physicist Robert Matthews suggested in 2001,^[158] is one way to avoid the problem that the prior probability is rarely known.

Encouraging larger sample sizes

To improve the quality of replications, larger sample sizes than those used in the original study are often needed.^[159] Larger sample sizes are needed because estimates of effect sizes in published work are often exaggerated due to publication bias and large sampling variability associated with small sample sizes in an original study.^{[160][161][162]} Further, using significance thresholds usually leads to inflated effects, because particularly with small sample sizes, only the largest effects will become significant.^[163]

Replication efforts

Funding

In July 2016, the Netherlands Organisation for Scientific Research made €3 million available for replication studies. The funding is for replication based on reanalysis of existing data and replication by collecting and analysing new data. Funding is available in the areas of social sciences, health research and healthcare innovation.^[164]

In 2013, the Laura and John Arnold Foundation funded the launch of The Center for Open Science with a \$5.25 million grant. By 2017, it provided an additional \$10 million in funding.^[165] It also funded the launch of the Meta-Research Innovation Center at Stanford at Stanford University run by Ioannidis and medical scientist Steven Goodman to study ways to improve scientific research.^[165] It also provided funding for the AllTrials initiative led in part by medical scientist Ben Goldacre.^[165]

Emphasis in post-secondary education

Based on coursework in experimental methods at MIT, Stanford, and the University of

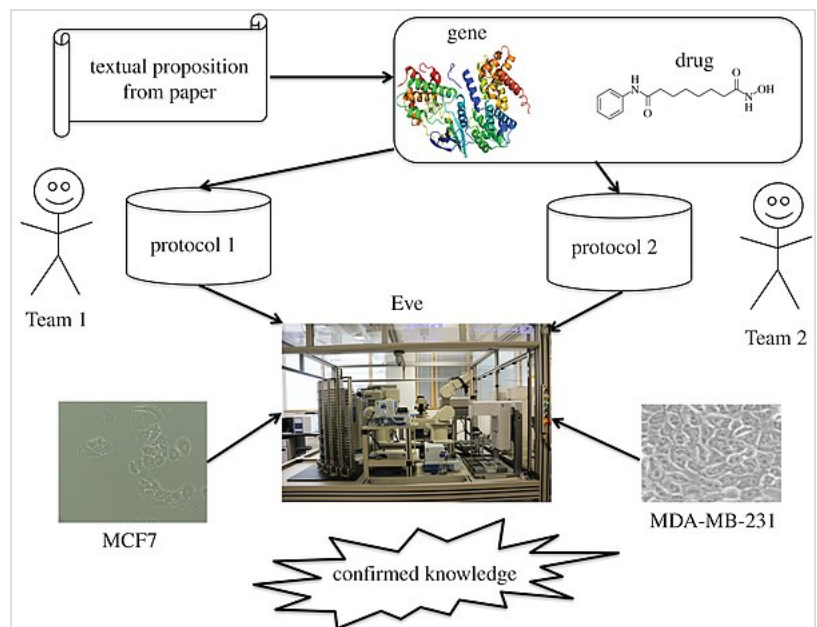
Washington, it has been suggested that methods courses in psychology and other fields should emphasize replication attempts rather than original studies.^{[166][167][168]} Such an approach would help students learn scientific methodology and provide numerous independent replications of meaningful scientific findings that would test the replicability of scientific findings. Some have recommended that graduate students should be required to publish a high-quality replication attempt on a topic related to their doctoral research prior to graduation.^[169]

Final year thesis

Some institutions require undergraduate students to submit a final year thesis that consists of an original piece of research. Daniel Quintana, a psychologist at the University of Oslo in Norway, has recommended that students should be encouraged to perform replication studies in thesis projects, as well as being taught about open science.^[170]

Semi-automated

Researchers demonstrated a way of semi-automated testing for reproducibility: statements about experimental results were extracted from, as of 2022 non-semantic, gene expression cancer research papers and subsequently reproduced via robot scientist "Eve".^{[171][172]} Problems of this approach include that it may not be feasible for many areas of research and that sufficient experimental data may not get extracted from some or many papers even if available.



"The overall process of testing the reproducibility and robustness of the cancer biology literature by robot. First, text mining is used to extract statements about the effect of drugs on gene expression in breast cancer. Then two different teams semi-automatically tested these statements using two different protocols, and two different cell lines (MCF7 and MDA-MB-231) using the laboratory automation system Eve."

Involving original authors

Psychologist Daniel Kahneman argued that, in psychology, the original authors should be involved in the replication effort because the published methods are often too vague.^{[173][174]} Others, such as psychologist Andrew Wilson, disagree, arguing that the original authors should write down the methods in detail.^[173] An investigation of replication rates in psychology in 2012 indicated higher success rates of replication in replication studies when there was author overlap with the original authors of a study^[175] (91.7% successful replication rates in studies with author overlap compared to 64.6% successful replication rates without author overlap).

Big team science

The replication crisis has led to the formation and development of various large-scale and

collaborative communities to pool their resources to address a single question across cultures, countries and disciplines.^[176] The focus is on replication, to ensure that the effect generalizes beyond a specific culture and investigate whether the effect is replicable and genuine.^[177] This allows interdisciplinary internal reviews, multiple perspectives, uniform protocols across labs, and recruiting larger and more diverse samples.^[177] Researchers can collaborate by coordinating data collection or fund data collection by researchers who may not have access to the funds, allowing larger sample sizes and increasing the robustness of the conclusions.

Broader changes to scientific approach

Emphasize triangulation, not just replication

Psychologist Marcus R. Munafò and Epidemiologist George Davey Smith argue, in a piece published by *Nature*, that research should emphasize triangulation, not just replication, to protect against flawed ideas. They claim that,

replication alone will get us only so far (and) might actually make matters worse ... [Triangulation] is the strategic use of multiple approaches to address one question. Each approach has its own unrelated assumptions, strengths and weaknesses. Results that agree across different methodologies are less likely to be artefacts. ... Maybe one reason replication has captured so much interest is the often-repeated idea that falsification is at the heart of the scientific enterprise. This idea was popularized by Karl Popper's 1950s maxim that theories can never be proved, only falsified. Yet an overemphasis on repeating experiments could provide an unfounded sense of certainty about findings that rely on a single approach. ... philosophers of science have moved on since Popper. Better descriptions of how scientists actually work include what epistemologist Peter Lipton called in 1991 "inference to the best explanation".^[178]

Complex systems paradigm

The dominant scientific and statistical model of causation is the linear model.^[179] The linear model assumes that mental variables are stable properties which are independent of each other. In other words, these variables are not expected to influence each other. Instead, the model assumes that the variables will have an independent, linear effect on observable outcomes.^[179]

Social scientists Sebastian Wallot and Damian Kelty-Stephen argue that the linear model is not always appropriate.^[179] An alternative is the complex system model which assumes that mental variables are interdependent. These variables are not assumed to be stable, rather they will interact and adapt to each specific context.^[179] They argue that the complex system model is often more appropriate in psychology, and that the use of the linear model when the complex system model is more appropriate will result in failed replications.^[179]

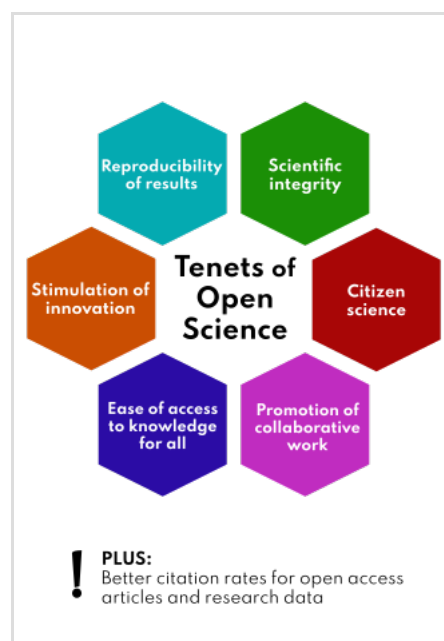
...psychology may be hoping for replications in the very measurements and under the very conditions where a growing body of psychological evidence explicitly discourages predicting replication. Failures to replicate may be plainly baked into the potentially incomplete, but broadly sweeping failure of human behavior to conform to the standard of independen[ce] ...^[179]

Replication should seek to revise theories

Replication is fundamental for scientific progress to confirm original findings. However, replication alone is not sufficient to resolve the replication crisis. Replication efforts should seek not just to support or question the original findings, but also to replace them with revised, stronger theories with greater explanatory power. This approach therefore involves pruning existing theories, comparing all the alternative theories, and making replication efforts more generative and engaged in theory-building.^{[180][181]} However, replication alone is not enough, it is important to assess the extent that results generalise across geographical, historical and social contexts is important for several scientific fields, especially practitioners and policy makers to make analyses in order to guide important strategic decisions. Reproducible and replicable findings was the best predictor of generalisability beyond historical and geographical contexts, indicating that for social sciences, results from a certain time period and place can meaningfully drive as to what is universally present in individuals.^[182]

Open science

Open data, open source software and open source hardware all are critical to enabling reproducibility in the sense of validation of the original data analysis. The use of proprietary software, the lack of the publication of analysis software and the lack of open data prevents the replication of studies. Unless software used in research is open source, reproducing results with different software and hardware configurations is impossible.^[183] CERN has both Open Data and CERN Analysis Preservation projects for storing data, all relevant information, and all software and tools needed to preserve an analysis at the large experiments of the LHC. Aside from all software and data, preserved analysis assets include metadata that enable understanding of the analysis workflow, related software, systematic uncertainties, statistics procedures and meaningful ways to search for the analysis, as well as references to publications and to backup material.^[184] CERN software is open source and available for use outside of particle physics and there is some guidance provided to other fields on the broad approaches and strategies used for open science in contemporary particle physics.^[185]



Tenets of open science

Online repositories where data, protocols, and findings can be stored and evaluated by the public seek to improve the integrity and reproducibility of research. Examples of such repositories include the Open Science Framework, Registry of Research Data Repositories, and Psychfiledrawer.org. Sites like Open Science Framework offer badges for using open science practices in an effort to incentivize scientists. However, there have been concerns that those who are most likely to provide their data and code for analyses are the researchers that are likely the most sophisticated.^[186] Ioannidis suggested that "the paradox may arise that the most meticulous and sophisticated and method-savvy and careful researchers may become more susceptible to criticism and reputation attacks by reanalyzers who hunt for errors, no matter how negligible these errors are".^[186]

See also

- [Decline effect](#)
- [Estimation statistics](#)
- [Exploratory data analysis](#)
- [p-value](#)
- [Student's t-test](#)
- [Analysis of variance](#)
- [Misuse of statistics](#)
- [Data dredging](#)
- [Invalid science](#)
- [Falsifiability](#)
- [Extension neglect](#)
- [Base rate fallacy](#)
- [Selection bias](#)
- [Sampling bias](#)
- [Observer bias](#)
- [Naturalism](#)
- [Problem of induction](#)
- [Correlation does not imply causation](#)
- [Black swan theory](#)
- [Uniformitarianism](#)

Notes

- a. According to the *APA Dictionary of Psychology*, confirmation bias is "the tendency to gather evidence that confirms preexisting expectations, typically by emphasizing or pursuing supporting evidence while dismissing or failing to seek contradictory evidence".^[105]
- b. In the context of null-hypothesis significance testing, results that are not statistically significant
- c. Data dredging, also known as *p-hacking* or *p-fishing*, is misuse of data, through myriad techniques, to find support for hypotheses that the data is inadequate for.^[112]
- d. Selective reporting is also known as *partial publication*. Reporting is an opportunity to disclose all of the researcher degrees of freedom used or exploited. Selective reporting is a failure to report relevant details or choices, such as some independent and dependent variables, missing data, data exclusions, and outlier exclusions.^[110]
- e. HARKing, also known as *post-hoc storytelling*, is when an exploratory analysis is framed as a confirmatory analysis. It involves changing a hypothesis after research has been done, so that the new hypothesis is able to be confirmed by the results of the experiment.^[110]
- f. The authors make an example whereby assuming that the true mean correlation reflecting an effect is 0.2 and the standard deviation of the distribution of effects is also 0.2, a replication study will have a 62% probability of finding either a medium-to-large true effect ($r > 0.3$) or a negligible true effect ($r < 0.1$).
- g. 0.412 against 0.389 in units of standardized mean differences (SMD).
- h. The main DV used was the subjective binary rating (i.e replicated/ not replicated) used in the original study by OSC. The authors also measured correlations with other measures of reproducibility (e.g. Confidence intervals) and found nearly-equal correlations between

context-sensitivity and replication success

- i. The independent effect of context-sensitivity could be observed both in a multiple logistic regression and in a hierarchical regression model. In the latter case, context-sensitivity was included in step 2 of the hierarchy and change in the coefficient of multiple determination turned out to be significant
- j. Following Bird's argument this percentage is obtained by calculating the False Positive Report Probability (FPRP) as follows.
 - $FPRP = \text{Number of false positives} / \text{Number of total positives}$
 - $\text{Number of false positives} = \text{Probability of obtaining a false positive} \times \text{Number of negative tests}$
 - $\text{Number of true positives} = \text{Probability of obtaining a true positive} \times \text{Number of positive tests}$

Assuming:

- Number of tests = 1000
- Proportion of true hypotheses $p = 0.10$
- Probability of obtaining a false positive $a = 0.05$
- Probability of obtaining a true positive $1 - B = 0.8$

Then $FPRP = (0.05 \times 900) / (0.05 \times 900 + 0.8 \times 100) = 0.36$

References

1. Ioannidis JP (August 2005). "Why most published research findings are false" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1182327>). *PLOS Medicine*. **2** (8): e124. doi:10.1371/journal.pmed.0020124 (<https://doi.org/10.1371%2Fjournal.pmed.0020124>). PMC 1182327 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1182327>). PMID 16060722 (<https://pubmed.ncbi.nlm.nih.gov/16060722>).
2. John S (8 December 2017). *Scientific Method*. New York, NY: Routledge. doi:10.4324/9781315100708 (<https://doi.org/10.4324%2F9781315100708>). ISBN 978-1-315-10070-8. S2CID 201781341 (<https://api.semanticscholar.org/CorpusID:201781341>).
3. Lehrer J (13 December 2010). "The Truth Wears Off" (<https://www.newyorker.com/magazine/2010/12/13/the-truth-wears-off>). *The New Yorker*. Retrieved 2020-01-30.
4. Marcus G (1 May 2013). "The Crisis in Social Psychology That Isn't" (<https://www.newyorker.com/tech/annals-of-technology/the-crisis-in-social-psychology-that-isnt>). *The New Yorker*. Retrieved 2020-01-30.
5. Baker M (May 2016). "1,500 scientists lift the lid on reproducibility" (<https://doi.org/10.1038%2F533452a>). *Nature* (News Feature). Springer Nature. **533** (7604): 452–454. Bibcode:2016Natur.533..452B (<https://ui.adsabs.harvard.edu/abs/2016Natur.533..452B>). doi:10.1038/533452a (<https://doi.org/10.1038%2F533452a>). PMID 27225100 (<https://pubmed.ncbi.nlm.nih.gov/27225100>). S2CID 4460617 (<https://api.semanticscholar.org/CorpusID:4460617>). (**Erratum: [1]** (<https://www.nature.com/articles/533452a#change-history>))
6. Pashler H, Harris CR (November 2012). "Is the Replicability Crisis Overblown? Three Arguments Examined". *Perspectives on Psychological Science*. **7** (6): 531–536. doi:10.1177/1745691612463401 (<https://doi.org/10.1177%2F1745691612463401>). PMID 26168109 (<https://pubmed.ncbi.nlm.nih.gov/26168109>). S2CID 1342421 (<https://api.semanticscholar.org/CorpusID:1342421>).

7. Fidler F, Wilcox J (2018). "Reproducibility of Scientific Results" (<https://plato.stanford.edu/entries/scientific-reproducibility/#MetaScieEstaMoniEvalReprCris>). *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Retrieved 2019-05-19.
8. Moonesinghe R, Khoury MJ, Janssens AC (February 2007). "Most published research findings are false-but a little replication goes a long way" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1808082>). *PLOS Medicine*. **4** (2): e28. doi:10.1371/journal.pmed.0040028 (<https://doi.org/10.1371%2Fjournal.pmed.0040028>). PMC 1808082 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1808082>). PMID 17326704 (<https://pubmed.ncbi.nlm.nih.gov/17326704>).
9. Simons DJ (January 2014). "The Value of Direct Replication". *Perspectives on Psychological Science*. **9** (1): 76–80. doi:10.1177/1745691613514755 (<https://doi.org/10.1177%2F1745691613514755>). PMID 26173243 (<https://pubmed.ncbi.nlm.nih.gov/26173243>). S2CID 1149441 (<https://api.semanticscholar.org/CorpusID:1149441>).
10. Schmidt S (2009). "Shall we Really do it Again? The Powerful Concept of Replication is Neglected in the Social Sciences". *Review of General Psychology*. SAGE Publications. **13** (2): 90–100. doi:10.1037/a0015108 (<https://doi.org/10.1037%2Fa0015108>). ISSN 1089-2680 (<https://www.worldcat.org/issn/1089-2680>). S2CID 143855611 (<https://api.semanticscholar.org/CorpusID:143855611>).
11. Open Science Collaboration (August 2015). "PSYCHOLOGY. Estimating the reproducibility of psychological science" (https://ink.library.smu.edu.sg/lkcsb_research/5257). *Science*. **349** (6251): aac4716. doi:10.1126/science.aac4716 (<https://doi.org/10.1126%2Fscience.aac4716>). hdl:10722/230596 (<https://hdl.handle.net/10722%2F230596>). PMID 26315443 (<https://pubmed.ncbi.nlm.nih.gov/26315443>). S2CID 218065162 (<https://api.semanticscholar.org/CorpusID:218065162>).
12. Duvendack M, Palmer-Jones R, Reed RW (May 2017). "What Is Meant by "Replication" and Why Does It Encounter Resistance in Economics?" (<https://www.aeaweb.org/articles?id=10.1257/aer.p20171031>). *American Economic Review*. **107** (5): 46–51. doi:10.1257/aer.p20171031 (<https://doi.org/10.1257%2Faer.p20171031>). ISSN 0002-8282 (<https://www.worldcat.org/issn/0002-8282>).
13. Shrout PE, Rodgers JL (January 2018). "Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis". *Annual Review of Psychology*. Annual Reviews. **69** (1): 487–510. doi:10.1146/annurev-psych-122216-011845 (<https://doi.org/10.1146%2Fannurev-psych-122216-011845>). PMID 29300688 (<https://pubmed.ncbi.nlm.nih.gov/29300688>). S2CID 19593610 (<https://api.semanticscholar.org/CorpusID:19593610>).
14. Stanley TD, Carter EC, Doucouliagos H (December 2018). "What meta-analyses reveal about the replicability of psychological research". *Psychological Bulletin*. **144** (12): 1325–1346. doi:10.1037/bul0000169 (<https://doi.org/10.1037%2Fbul0000169>). PMID 30321017 (<https://pubmed.ncbi.nlm.nih.gov/30321017>). S2CID 51951232 (<https://api.semanticscholar.org/CorpusID:51951232>).
15. Meyer C, Chabris C (31 July 2014). "Why Psychologists' Food Fight Matters" (http://www.slate.com/articles/health_and_science/science/2014/07/replication_controversy_in_psychology_bullying_file_drawer_effect_blog_posts.single.html). *Slate*.
16. Aschwanden C (19 August 2015). "Science Isn't Broken" (<https://fivethirtyeight.com/features/science-isnt-broken/>). *FiveThirtyEight*. Retrieved 2020-01-30.
17. Aschwanden C (27 August 2015). "Psychology Is Starting To Deal With Its Replication Problem" (<https://fivethirtyeight.com/features/psychology-is-starting-to-deal-with-its-replication-problem/>). *FiveThirtyEight*. Retrieved 2020-01-30.
18. Etchells P (28 May 2014). "Psychology's replication drive: it's not about you" (<https://www.theguardian.com/science/head-quarters/2014/may/28/psychology-replication-drive-methods-bullying>). *The Guardian*.

19. Wagenmakers EJ, Wetzels R, Borsboom D, van der Maas HL, Kievit RA (November 2012). "An Agenda for Purely Confirmatory Research". *Perspectives on Psychological Science*. **7** (6): 632–638. doi:10.1177/1745691612463078 (<https://doi.org/10.1177%2F1745691612463078>). PMID 26168122 (<https://pubmed.ncbi.nlm.nih.gov/26168122>). S2CID 5096417 (<https://api.semanticscholar.org/CorpusID:5096417>).
20. Ioannidis JP (November 2012). "Why Science Is Not Necessarily Self-Correcting". *Perspectives on Psychological Science*. **7** (6): 645–654. doi:10.1177/1745691612464056 (<https://doi.org/10.1177%2F1745691612464056>). PMID 26168125 (<https://pubmed.ncbi.nlm.nih.gov/26168125>). S2CID 11798785 (<https://api.semanticscholar.org/CorpusID:11798785>).
21. Pashler H, Harris CR (November 2012). "Is the Replicability Crisis Overblown? Three Arguments Examined". *Perspectives on Psychological Science*. **7** (6): 531–536. doi:10.1177/1745691612463401 (<https://doi.org/10.1177%2F1745691612463401>). PMID 26168109 (<https://pubmed.ncbi.nlm.nih.gov/26168109>). S2CID 1342421 (<https://api.semanticscholar.org/CorpusID:1342421>).
22. Meehl PE (1967). "Theory-Testing in Psychology and Physics: A Methodological Paradox" (<https://www.jstor.org/stable/186099>). *Philosophy of Science*. **34** (2): 103–115. doi:10.1086/288135 (<https://doi.org/10.1086%2F288135>). ISSN 0031-8248 (<https://www.worldcat.org/issn/0031-8248>). JSTOR 186099 (<https://www.jstor.org/stable/186099>). S2CID 96422880 (<https://api.semanticscholar.org/CorpusID:96422880>).
23. Romero F (November 2019). "Philosophy of science and the replicability crisis" (<https://doi.org/10.1111%2Fphc3.12633>). *Philosophy Compass*. **14** (11). doi:10.1111/phc3.12633 (<https://doi.org/10.1111%2Fphc3.12633>). ISSN 1747-9991 (<https://www.worldcat.org/issn/1747-9991>). S2CID 202261836 (<https://api.semanticscholar.org/CorpusID:202261836>).
24. Bargh JA, Chen M, Burrows L (August 1996). "Automaticity of social behavior: direct effects of trait construct and stereotype-activation on action". *Journal of Personality and Social Psychology*. **71** (2): 230–244. doi:10.1037/0022-3514.71.2.230 (<https://doi.org/10.1037%2F0022-3514.71.2.230>). PMID 8765481 (<https://pubmed.ncbi.nlm.nih.gov/8765481>). S2CID 6654763 (<https://api.semanticscholar.org/CorpusID:6654763>).
25. Doyen S, Klein O, Pichon CL, Cleeremans A (18 January 2012). Lauwereyns J (ed.). "Behavioral priming: it's all in the mind, but whose mind?" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3261136>). *PLOS ONE*. **7** (1): e29081. Bibcode:2012PLoSO...729081D (<https://ui.adsabs.harvard.edu/abs/2012PLoSO...729081D>). doi:10.1371/journal.pone.0029081 (<https://doi.org/10.1371%2Fjournal.pone.0029081>). PMC 3261136 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3261136>). PMID 22279526 (<https://pubmed.ncbi.nlm.nih.gov/22279526>).
26. Yong E (10 March 2012). "A failed replication draws a scathing personal attack from a psychology professor" (<https://web.archive.org/web/20210225023644/https://www.nationalgeographic.com/science/article/failed-replication-bargh-psychology-study-doyen>). *National Geographic*. Archived from the original (<https://www.nationalgeographic.com/science/article/failed-replication-bargh-psychology-study-doyen>) on 2021-02-25. Retrieved 2023-07-04.
27. Pashler H, Coburn N, Harris CR (29 August 2012). "Priming of social distance? Failure to replicate effects on social and food judgments" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3430642>). *PLOS ONE*. **7** (8): e42510. Bibcode:2012PLoSO...742510P (<https://ui.adsabs.harvard.edu/abs/2012PLoSO...742510P>). doi:10.1371/journal.pone.0042510 (<https://doi.org/10.1371%2Fjournal.pone.0042510>). PMC 3430642 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3430642>). PMID 22952597 (<https://pubmed.ncbi.nlm.nih.gov/22952597>).
28. Harris CR, Coburn N, Rohrer D, Pashler H (16 August 2013). "Two failures to replicate high-performance-goal priming effects" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3745413>). *PLOS ONE*. **8** (8): e72467. Bibcode:2013PLoSO...872467H (<https://ui.adsabs.harvard.edu/abs/2013PLoSO...872467H>). doi:10.1371/journal.pone.0072467 (<https://doi.org/10.1371%2Fjournal.pone.0072467>). PMC 3745413 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3745413>). PMID 23977304 (<https://pubmed.ncbi.nlm.nih.gov/23977304>).

29. Shanks DR, Newell BR, Lee EH, Balakrishnan D, Ekelund L, Cenac Z, et al. (24 April 2013). "Priming intelligent behavior: an elusive phenomenon" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3634790>). *PLOS ONE*. **8** (4): e56515. Bibcode:2013PLoSO...856515S (<https://ui.adsabs.harvard.edu/abs/2013PLoSO...856515S>). doi:10.1371/journal.pone.0056515 (<https://doi.org/10.1371%2Fjournal.pone.0056515>). PMC 3634790 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3634790>). PMID 23637732 (<https://pubmed.ncbi.nlm.nih.gov/23637732>).
30. Klein RA, Ratliff KA, Vianello M, Adams RB, Bahník Š, Bernstein MJ, Bocian K, Brandt MJ, Brooks B, Brumbaugh CC, Cemalcilar Z, Chandler J, Cheong W, Davis WE, Devos T (May 2014). "Investigating Variation in Replicability" (<https://doi.org/10.1027%2F1864-9335%2Fa000178>). *Social Psychology*. **45** (3): 142–152. doi:10.1027/1864-9335/a000178 (<https://doi.org/10.1027%2F1864-9335%2Fa000178>). hdl:2066/131506 (<https://hdl.handle.net/2066%2F131506>). ISSN 1864-9335 (<https://www.worldcat.org/issn/1864-9335>).
31. Bem DJ (March 2011). "Feeling the future: experimental evidence for anomalous retroactive influences on cognition and affect". *Journal of Personality and Social Psychology*. **100** (3): 407–425. doi:10.1037/a0021524 (<https://doi.org/10.1037%2Fa0021524>). PMID 21280961 (<https://pubmed.ncbi.nlm.nih.gov/21280961>). S2CID 1961013 (<https://api.semanticscholar.org/CorpusID:1961013>).
32. Wagenmakers EJ, Wetzels R, Borsboom D, van der Maas HL (March 2011). "Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011)". *Journal of Personality and Social Psychology*. **100** (3): 426–432. doi:10.1037/a0022790 (<https://doi.org/10.1037%2Fa0022790>). PMID 21280965 (<https://pubmed.ncbi.nlm.nih.gov/21280965>).
33. Galak J, LeBoeuf RA, Nelson LD, Simmons JP (December 2012). "Correcting the past: failures to replicate ψ " (<https://zenodo.org/record/889659>). *Journal of Personality and Social Psychology*. **103** (6): 933–948. doi:10.1037/a0029709 (<https://doi.org/10.1037%2Fa0029709>). PMID 22924750 (<https://pubmed.ncbi.nlm.nih.gov/22924750>).
34. Begley CG, Ellis LM (March 2012). "Drug development: Raise standards for preclinical cancer research" (<https://doi.org/10.1038%2F483531a>). *Nature*. **483** (7391): 531–533. Bibcode:2012Natur.483..531B (<https://ui.adsabs.harvard.edu/abs/2012Natur.483..531B>). doi:10.1038/483531a (<https://doi.org/10.1038%2F483531a>). PMID 22460880 (<https://pubmed.ncbi.nlm.nih.gov/22460880>). S2CID 4326966 (<https://api.semanticscholar.org/CorpusID:4326966>).
35. Prinz F, Schlange T, Asadullah K (August 2011). "Believe it or not: how much can we rely on published data on potential drug targets?" (<https://doi.org/10.1038%2Fnrd3439-c1>). *Nature Reviews. Drug Discovery*. **10** (9): 712. doi:10.1038/nrd3439-c1 (<https://doi.org/10.1038%2Fnrd3439-c1>). PMID 21892149 (<https://pubmed.ncbi.nlm.nih.gov/21892149>). S2CID 16180896 (<https://api.semanticscholar.org/CorpusID:16180896>).
36. Ioannidis JP (September 2008). "Why most discovered true associations are inflated" (<https://doi.org/10.1097%2FEDE.0b013e31818131e7>). *Epidemiology*. **19** (5): 640–648. doi:10.1097/EDE.0b013e31818131e7 (<https://doi.org/10.1097%2FEDE.0b013e31818131e7>). PMID 18633328 (<https://pubmed.ncbi.nlm.nih.gov/18633328>). S2CID 15440816 (<https://api.semanticscholar.org/CorpusID:15440816>).
37. John LK, Loewenstein G, Prelec D (May 2012). "Measuring the prevalence of questionable research practices with incentives for truth telling". *Psychological Science*. **23** (5): 524–532. doi:10.1177/0956797611430953 (<https://doi.org/10.1177%2F0956797611430953>). PMID 22508865 (<https://pubmed.ncbi.nlm.nih.gov/22508865>). S2CID 8400625 (<https://api.semanticscholar.org/CorpusID:8400625>).
38. Simmons JP, Nelson LD, Simonsohn U (November 2011). "False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant". *Psychological Science*. **22** (11): 1359–1366. doi:10.1177/0956797611417632 (<https://doi.org/10.1177%2F0956797611417632>). PMID 22006061 (<https://pubmed.ncbi.nlm.nih.gov/22006061>). S2CID 13802986 (<https://api.semanticscholar.org/CorpusID:13802986>).

39. Pashler H, Wagenmakers EJ (November 2012). "Editors' Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence?". *Perspectives on Psychological Science*. **7** (6): 528–530. doi:10.1177/1745691612465253 (<https://doi.org/10.1177%2F1745691612465253>). PMID 26168108 (<https://pubmed.ncbi.nlm.nih.gov/26168108>). S2CID 26361121 (<https://api.semanticscholar.org/CorpusID:26361121>).
40. Ahlgren, Andrew (April 1969). "A modest proposal for encouraging replication" (<http://doi.apa.org/getdoi.cfm?doi=10.1037/h0037798>). *American Psychologist*. **24** (4): 471. doi:10.1037/h0037798 (<https://doi.org/10.1037%2Fh0037798>). ISSN 1935-990X (<https://www.worldcat.org/issn/1935-990X>).
41. Smith, Nathaniel C. (October 1970). "Replication studies: A neglected aspect of psychological research" (<http://doi.apa.org/getdoi.cfm?doi=10.1037/h0029774>). *American Psychologist*. **25** (10): 970–975. doi:10.1037/h0029774 (<https://doi.org/10.1037%2Fh0029774>). ISSN 1935-990X (<https://www.worldcat.org/issn/1935-990X>).
42. Neuliep, J. W.; Crandall, R. (1993). "Reviewer bias against replication research" (<https://www.proquest.com/docview/1292304227>). *Journal of Social Behavior and Personality*. **8** (6): 21–29. ProQuest 1292304227 (<https://search.proquest.com/docview/1292304227>) – via ProQuest.
43. Neuliep, J. W.; Crandall, R. (1990). "Editorial bias against replication research" (<https://www.proquest.com/openview/ecff1f2487c66017c85a0ead234c9168/1?pq-origsite=gscholar&cbI=1819046>). *Journal of Social Behavior and Personality*. **5** (4): 85–90 – via ProQuest.
44. Lewis-Kraus, Gideon (30 September 2023). "They Studied Dishonesty. Was Their Work a Lie?" (<https://www.newyorker.com/magazine/2023/10/09/they-studied-dishonesty-was-their-work-a-lie>). *The New Yorker*. ISSN 0028-792X (<https://www.worldcat.org/issn/0028-792X>). Retrieved 2023-10-01.
45. Subbaraman, Nidhi (24 September 2023). "The Band of Debunkers Busting Bad Scientists" (<https://www.wsj.com/science/data-colada-debunk-stanford-president-research-14664f3>). *Wall Street Journal*. Archived (<https://archive.today/20230924094046/https://www.wsj.com/science/data-colada-debunk-stanford-president-research-14664f3>) from the original on 2023-09-24. Retrieved 2023-10-08.
46. "APA PsycNet" (<https://psycnet.apa.org/record/2013-25331-001>). *psycnet.apa.org*. Retrieved 2023-10-08.
47. Spellman, Barbara A. (November 2015). "A Short (Personal) Future History of Revolution 2.0" (<https://doi.org/10.1177%2F1745691615609918>). *Perspectives on Psychological Science*. **10** (6): 886–899. doi:10.1177/1745691615609918 (<https://doi.org/10.1177%2F1745691615609918>). ISSN 1745-6916 (<https://www.worldcat.org/issn/1745-6916>). PMID 26581743 (<https://pubmed.ncbi.nlm.nih.gov/26581743>). S2CID 206778431 (<https://api.semanticscholar.org/CorpusID:206778431>).
48. Greenwald, Anthony G., ed. (January 1976). "An editorial" (<http://doi.apa.org/getdoi.cfm?doi=10.1037/h0078635>). *Journal of Personality and Social Psychology*. **33** (1): 1–7. doi:10.1037/h0078635 (<https://doi.org/10.1037%2Fh0078635>). ISSN 1939-1315 (<https://www.worldcat.org/issn/1939-1315>).
49. Sterling, Theodore D. (1959). "Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance--Or Vice Versa" (<https://www.jstor.org/stable/2282137>). *Journal of the American Statistical Association*. **54** (285): 30–34. doi:10.2307/2282137 (<https://doi.org/10.2307%2F2282137>). ISSN 0162-1459 (<https://www.worldcat.org/issn/0162-1459>). JSTOR 2282137 (<https://www.jstor.org/stable/2282137>).
50. Mills, J. L. (14 October 1993). "Data torturing" (<https://pubmed.ncbi.nlm.nih.gov/8166792/>). *The New England Journal of Medicine*. **329** (16): 1196–1199. doi:10.1056/NEJM199310143291613 (<https://doi.org/10.1056%2FNEJM199310143291613>). ISSN 0028-4793 (<https://www.worldcat.org/issn/0028-4793>). PMID 8166792 (<https://pubmed.ncbi.nlm.nih.gov/8166792/>).

51. Rosenthal, Robert (May 1979). "The file drawer problem and tolerance for null results" (<http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-2909.86.3.638>). *Psychological Bulletin*. **86** (3): 638–641. doi:10.1037/0033-2909.86.3.638 (<https://doi.org/10.1037%2F0033-2909.86.3.638>). ISSN 1939-1455 (<https://www.worldcat.org/issn/1939-1455>). S2CID 36070395 (<https://api.semanticscholar.org/CorpusID:36070395>).
52. Cohen, J. (September 1962). "The statistical power of abnormal-social psychological research: a review" (<https://pubmed.ncbi.nlm.nih.gov/13880271/>). *Journal of Abnormal and Social Psychology*. **65**: 145–153. doi:10.1037/h0045186 (<https://doi.org/10.1037%2Fh0045186>). ISSN 0096-851X (<https://www.worldcat.org/issn/0096-851X>). PMID 13880271 (<https://pubmed.ncbi.nlm.nih.gov/13880271/>).
53. Sedlmeier, Peter; Gigerenzer, Gerd (March 1989). "Do studies of statistical power have an effect on the power of studies?" (<http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-2909.105.2.309>). *Psychological Bulletin*. **105** (2): 309–316. doi:10.1037/0033-2909.105.2.309 (<https://doi.org/10.1037%2F0033-2909.105.2.309>). hdl:21.1116/0000-0000-B883-C (<https://hdl.handle.net/21.1116%2F0000-0000-B883-C>). ISSN 1939-1455 (<https://www.worldcat.org/issn/1939-1455>).
54. Achenbach J. "No, science's reproducibility problem is not limited to psychology" (<https://www.washingtonpost.com/news/speaking-of-science/wp/2015/08/28/no-sciences-reproducibility-problem-is-not-limited-to-psychology/>). *The Washington Post*. Retrieved 2015-09-10.
55. Wiggins BJ, Christopherson C (2019). "The replication crisis in psychology: An overview for theoretical and philosophical psychology" (<http://doi.apa.org/getdoi.cfm?doi=10.1037/teo0000137>). *Journal of Theoretical and Philosophical Psychology*. **39** (4): 202–217. doi:10.1037/teo0000137 (<https://doi.org/10.1037%2Fteo0000137>). ISSN 2151-3341 (<https://www.worldcat.org/issn/2151-3341>). S2CID 210567289 (<https://api.semanticscholar.org/CorpusID:210567289>).
56. Bartlett T (30 January 2013). "Power of Suggestion" (<http://chronicle.com/article/Power-of-Suggestion/136907/>). *The Chronicle of Higher Education*.
57. Dominus S (18 October 2017). "When the Revolution Came for Amy Cuddy" (<https://www.nytimes.com/2017/10/18/magazine/when-the-revolution-came-for-amy-cuddy.html>). *The New York Times*. ISSN 0362-4331 (<https://www.worldcat.org/issn/0362-4331>). Retrieved 2017-10-19.
58. Duncan LE, Keller MC (October 2011). "A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3222234>). *The American Journal of Psychiatry*. **168** (10): 1041–1049. doi:10.1176/appi.ajp.2011.11020191 (<https://doi.org/10.1176%2Fappi.ajp.2011.11020191>). PMC 3222234 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3222234>). PMID 21890791 (<https://pubmed.ncbi.nlm.nih.gov/21890791/>).
59. Leichsenring F, Abbass A, Hilsenroth MJ, Leweke F, Luyten P, Keefe JR, et al. (April 2017). "Biases in research: risk factors for non-replicability in psychotherapy and pharmacotherapy research" (<https://discovery.ucl.ac.uk/id/eprint/1532689/>). *Psychological Medicine*. **47** (6): 1000–1011. doi:10.1017/S003329171600324X (<https://doi.org/10.1017%2FS003329171600324X>). PMID 27955715 (<https://pubmed.ncbi.nlm.nih.gov/27955715/>). S2CID 1872762 (<https://api.semanticscholar.org/CorpusID:1872762>).
60. Hengartner MP (28 February 2018). "Raising Awareness for the Replication Crisis in Clinical Psychology by Focusing on Inconsistencies in Psychotherapy Research: How Much Can We Rely on Published Findings from Efficacy Trials?" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5835722>). *Frontiers in Psychology*. Frontiers Media. **9**: 256. doi:10.3389/fpsyg.2018.00256 (<https://doi.org/10.3389%2Ffpsyg.2018.00256>). PMC 5835722 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5835722>). PMID 29541051 (<https://pubmed.ncbi.nlm.nih.gov/29541051/>).

61. Frank MC, Bergelson E, Bergmann C, Cristia A, Floccia C, Gervain J, et al. (9 March 2017). "A Collaborative Approach to Infant Research: Promoting Reproducibility, Best Practices, and Theory-Building" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6879177>). *Infancy*. **22** (4): 421–435. doi:10.1111/infa.12182 (<https://doi.org/10.1111%2Finfa.12182>). hdl:10026.1/9942 (<https://hdl.handle.net/10026.1%2F9942>). PMC 6879177 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6879177>). PMID 31772509 (<https://pubmed.ncbi.nlm.nih.gov/31772509>).
62. Harris JR (2009) [1998]. *The Nurture Assumption: Why Children Turn Out the Way They Do* (2nd ed.). New York: Free Press. ISBN 978-1439101650.
63. Harris HR (2006). *No Two Alike: Human Nature and Human Individuality*. New York: W. W. Norton & Company. ISBN 978-0393329711.
64. Tyson C (14 August 2014). "Failure to Replicate" (<https://www.insidehighered.com/news/2014/08/14/almost-no-education-research-replicated-new-article-shows>). *Inside Higher Ed*. Retrieved 2018-12-19.
65. Makel MC, Plucker JA (1 August 2014). "Facts Are More Important Than Novelty: Replication in the Education Sciences" (<https://journals.sagepub.com/stoken/rbtf/w5mrNxPVD8zSg/full>). *Educational Researcher*. **43** (6): 304–316. doi:10.3102/0013189X14545513 (<https://doi.org/10.3102%2F0013189X14545513>). S2CID 145571836 (<https://api.semanticscholar.org/CorpusID:145571836>). Retrieved 2018-12-19.
66. Kirschner PA, Sweller J, Clark RE (2006). "Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching" (<https://www.researchgate.net/publication/27699659>). *Educational Psychologist*. Routledge. **41** (2): 75–86. doi:10.1207/s15326985ep4102_1 (https://doi.org/10.1207%2Fs15326985ep4102_1). S2CID 17067829 (<https://api.semanticscholar.org/CorpusID:17067829>).
67. Foundations for Success: The Final Report of the National Mathematics Advisory Panel (<https://www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf>) (PDF) (Report). United States Department of Education. 2008. pp. 45–46. Archived (<https://web.archive.org/web/20180118012046/https://www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf>) (PDF) from the original on 2018-01-18. Retrieved 2020-11-03.
68. Pashler H, McDaniel M, Rohrer D, Bjork R (December 2008). "Learning Styles: Concepts and Evidence" (<https://doi.org/10.1111%2Fj.1539-6053.2009.01038.x>). *Psychological Science in the Public Interest*. SAGE Publications. **9** (3): 105–119. doi:10.1111/j.1539-6053.2009.01038.x (<https://doi.org/10.1111%2Fj.1539-6053.2009.01038.x>). PMID 26162104 (<https://pubmed.ncbi.nlm.nih.gov/26162104>). S2CID 2112166 (<https://api.semanticscholar.org/CorpusID:2112166>).
69. Nosek BA, Cohoon J, Kidwell MC, Spies JR (2018) [2015]. "Summary of reproducibility rates and effect sizes for original and replication studies overall and by journal/discipline" (<https://mfr.osf.io/render?url=https://osf.io/jq7v6/?action=download%26mode=render>). *Estimating the Reproducibility of Psychological Science* (table). Reproducibility Project: Psychology. Retrieved 2019-10-16.
70. Roger A (27 August 2018). "The Science Behind Social Science Gets Shaken Up—Again" (<https://www.wired.com/story/social-science-reproducibility/>). *Wired*. Retrieved 2018-08-28.
71. Camerer CF, Dreber A, Holzmeister F, Ho TH, Huber J, Johannesson M, et al. (September 2018). "Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015" (<https://authors.library.caltech.edu/91063/>). *Nature Human Behaviour*. **2** (9): 637–644. doi:10.1038/s41562-018-0399-z (<https://doi.org/10.1038%2Fs41562-018-0399-z>). PMID 31346273 (<https://pubmed.ncbi.nlm.nih.gov/31346273>). S2CID 52098703 (<https://api.semanticscholar.org/CorpusID:52098703>).

72. Klein RA (2018). "Many Labs 2: Investigating Variation in Replicability Across Samples and Settings" (<https://doi.org/10.1177%2F2515245918810225>). *Advances in Methods and Practices in Psychological Science*. **1** (4): 443–490. doi:10.1177/2515245918810225 (<https://doi.org/10.1177%2F2515245918810225>). hdl:1854/LU-8637133 (<https://hdl.handle.net/1854%2FLU-8637133>).
73. Witkowski T (2019). "Is the glass half empty or half full? Latest results in the replication crisis in Psychology" (<https://web.archive.org/web/20200130130328/https://forbiddenpsychology.files.wordpress.com/2019/04/pdf-article.pdf>) (PDF). *Skeptical Inquirer*. Vol. 43, no. 2. pp. 5–6. Archived from the original (<https://forbiddenpsychology.files.wordpress.com/2019/04/pdf-article.pdf>) (PDF) on 2020-01-30.
74. Richtel M (16 March 2022). "Brain-Imaging Studies Hampered by Small Data Sets, Study Finds" (<https://www.nytimes.com/2022/03/16/science/brain-imaging-research.html>). *The New York Times*.
75. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, et al. (March 2022). "Reproducible brain-wide association studies require thousands of individuals" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8991999>). *Nature*. **603** (7902): 654–660. Bibcode:2022Natur.603..654M (<https://ui.adsabs.harvard.edu/abs/2022Natur.603..654M>). doi:10.1038/s41586-022-04492-9 (<https://doi.org/10.1038%2Fs41586-022-04492-9>). PMC 8991999 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8991999>). PMID 35296861 (<https://pubmed.ncbi.nlm.nih.gov/35296861>).
76. Ioannidis JP (July 2005). "Contradicted and initially stronger effects in highly cited clinical research". *JAMA*. **294** (2): 218–228. doi:10.1001/jama.294.2.218 (<https://doi.org/10.1001%2Fjama.294.2.218>). PMID 16014596 (<https://pubmed.ncbi.nlm.nih.gov/16014596>). S2CID 16749356 (<https://api.semanticscholar.org/CorpusID:16749356>).
77. Prinz F, Schlange T, Asadullah K (August 2011). "Believe it or not: how much can we rely on published data on potential drug targets?" (<https://doi.org/10.1038%2Fnrd3439-c1>). *Nature Reviews. Drug Discovery*. **10** (9): 712. doi:10.1038/nrd3439-c1 (<https://doi.org/10.1038%2Fnrd3439-c1>). PMID 21892149 (<https://pubmed.ncbi.nlm.nih.gov/21892149>).
78. Wheeling K (12 May 2016). "Big Pharma Reveals a Biomedical Replication Crisis" (<https://psmag.com/news/big-pharma-reveals-a-biomedical-replication-crisis>). *Pacific Standard*. Retrieved 2020-01-30. Updated on 14 June 2017
79. Begley CG, Ellis LM (March 2012). "Drug development: Raise standards for preclinical cancer research" (<https://doi.org/10.1038%2F483531a>). *Nature* (Comment article). **483** (7391): 531–533. Bibcode:2012Natur.483..531B (<https://ui.adsabs.harvard.edu/abs/2012Natur.483..531B>). doi:10.1038/483531a (<https://doi.org/10.1038%2F483531a>). PMID 22460880 (<https://pubmed.ncbi.nlm.nih.gov/22460880>). S2CID 4326966 (<https://api.semanticscholar.org/CorpusID:4326966>). (Erratum: doi:10.1038/485041e (<https://doi.org/10.1038%2F485041e>))
80. Haele T (7 December 2021). "Dozens of major cancer studies can't be replicated" (<https://www.sciencenews.org/article/cancer-biology-studies-research-replication-reproducibility>). *Science News*. Retrieved 2022-01-19.
81. "Reproducibility Project: Cancer Biology" (<https://www.cos.io/rpcb>). *www.cos.io*. Center for Open Science. Retrieved 2022-01-19.
82. Mobley A, Linder SK, Braeuer R, Ellis LM, Zwelling L (2013). Arakawa H (ed.). "A survey on data reproducibility in cancer research provides insights into our limited ability to translate findings from the laboratory to the clinic" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3655010>). *PLOS ONE*. **8** (5): e63221. Bibcode:2013PLoSO...863221M (<https://ui.adsabs.harvard.edu/abs/2013PLoSO...863221M>). doi:10.1371/journal.pone.0063221 (<https://doi.org/10.1371%2Fjournal.pone.0063221>). PMC 3655010 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3655010>). PMID 23691000 (<https://pubmed.ncbi.nlm.nih.gov/23691000>).

83. Van Noorden, Richard (18 July 2023). "Medicine is plagued by untrustworthy clinical trials. How many studies are faked or flawed?" (<https://doi.org/10.1038%2Fd41586-023-02299-w>). *Nature*. **619** (7970): 454–458. Bibcode:2023Natur.619..454V (<https://ui.adsabs.harvard.edu/abs/2023Natur.619..454V>). doi:10.1038/d41586-023-02299-w (<https://doi.org/10.1038%2Fd41586-023-02299-w>). PMID 37464079 (<https://pubmed.ncbi.nlm.nih.gov/37464079>).
84. Tsui AS (21 January 2022). "From Traditional Research to Responsible Research: The Necessity of Scientific Freedom and Scientific Responsibility for Better Societies" (<https://doi.org/10.1146%2Fannurev-orgpsych-062021-021303>). *Annual Review of Organizational Psychology and Organizational Behavior*. **9** (1): 1–32. doi:10.1146/annurev-orgpsych-062021-021303 (<https://doi.org/10.1146%2Fannurev-orgpsych-062021-021303>). ISSN 2327-0608 (<https://www.worldcat.org/issn/2327-0608>). S2CID 244238570 (<https://api.semanticscholar.org/CorpusID:244238570>).
85. Camerer CF, Dreber A, Forsell E, Ho TH, Huber J, Johannesson M, et al. (March 2016). "Evaluating replicability of laboratory experiments in economics" (<https://doi.org/10.1126%2Fscience.aaf0918>). *Science*. **351** (6280): 1433–1436. Bibcode:2016Sci...351.1433C (<https://ui.adsabs.harvard.edu/abs/2016Sci...351.1433C>). doi:10.1126/science.aaf0918 (<https://doi.org/10.1126%2Fscience.aaf0918>). PMID 26940865 (<https://pubmed.ncbi.nlm.nih.gov/26940865>).
86. Bohannon J (3 March 2016). "About 40% of economics experiments fail replication survey" (<https://www.science.org/content/article/about-40-economics-experiments-fail-replication-survey-rev2>). *Science*. doi:10.1126/science.aaf4141 (<https://doi.org/10.1126%2Fscience.aaf4141>). Retrieved 2017-10-25.
87. Goldfarb RS (1 December 1997). "Now you see it, now you don't: emerging contrary results in economics". *Journal of Economic Methodology*. **4** (2): 221–244. doi:10.1080/13501789700000016 (<https://doi.org/10.1080%2F13501789700000016>). ISSN 1350-178X (<https://www.worldcat.org/issn/1350-178X>).
88. Bergh DD, Sharp BM, Aguinis H, Li M (6 April 2017). "Is there a credibility crisis in strategic management research? Evidence on the reproducibility of study findings" (<https://doi.org/10.1177%2F1476127017701076>). *Strategic Organization*. **15** (3): 423–436. doi:10.1177/1476127017701076 (<https://doi.org/10.1177%2F1476127017701076>). ISSN 1476-1270 (<https://www.worldcat.org/issn/1476-1270>). S2CID 44024633 (<https://api.semanticscholar.org/CorpusID:44024633>).
89. Stagge JH, Rosenberg DE, Abdallah AM, Akbar H, Attallah NA, James R (February 2019). "Assessing data availability and research reproducibility in hydrology and water resources" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6390703>). *Scientific Data*. **6**: 190030. Bibcode:2019NatSD...690030S (<https://ui.adsabs.harvard.edu/abs/2019NatSD...690030S>). doi:10.1038/sdata.2019.30 (<https://doi.org/10.1038%2Fsdata.2019.30>). PMC 6390703 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6390703>). PMID 30806638 (<https://pubmed.ncbi.nlm.nih.gov/30806638>).
90. Nature Video (28 May 2016). "Is There a Reproducibility Crisis in Science?" (<https://www.scientificamerican.com/video/is-there-a-reproducibility-crisis-in-science/>). *Scientific American*. Retrieved 2019-08-15.
91. Fanelli, Daniele (2010). Enrico Scalas (ed.). "'Positive' Results Increase Down the Hierarchy of the Sciences" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2850928>). *PLOS ONE*. **5** (4): e10068. Bibcode:2010PLoS...510068F (<https://ui.adsabs.harvard.edu/abs/2010PLoSO...510068F>). doi:10.1371/journal.pone.0010068 (<https://doi.org/10.1371%2Fjournal.pone.0010068>). PMC 2850928 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2850928>). PMID 20383332 (<https://pubmed.ncbi.nlm.nih.gov/20383332>).

92. Allen C, Mehler DM (May 2019). "Open science challenges, benefits and tips in early career and beyond" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6513108>). *PLOS Biology*. Public Library of Science. **17** (5): e3000246. doi:10.1371/journal.pbio.3000246 (<https://doi.org/10.1371%2Fjournal.pbio.3000246>). PMC 6513108 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6513108>). PMID 31042704 (<https://pubmed.ncbi.nlm.nih.gov/31042704>).
93. Begley CG, Ioannidis JP (January 2015). "Reproducibility in science: improving the standard for basic and preclinical research" (<https://doi.org/10.1161%2FCIRCRESAHA.114.303819>). *Circulation Research*. **116** (1): 116–126. doi:10.1161/CIRCRESAHA.114.303819 (<https://doi.org/10.1161%2FCIRCRESAHA.114.303819>). PMID 25552691 (<https://pubmed.ncbi.nlm.nih.gov/25552691>). S2CID 3587510 (<https://api.semanticscholar.org/CorpusID:3587510>).
94. Price DJ (1963). *Little science big science* (<https://archive.org/details/littlesciencebig0000price>). Columbia University Press. p. 32. ISBN 9780231085625.
95. Siebert S, Machesky LM, Insall RH (September 2015). "Overflow in science and its implications for trust" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4563216>). *eLife*. **4**: e10825. doi:10.7554/eLife.10825 (<https://doi.org/10.7554%2FeLife.10825>). PMC 4563216 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4563216>). PMID 26365552 (<https://pubmed.ncbi.nlm.nih.gov/26365552>).
96. Della Briotta Parolo P, Pan RK, Ghosh R, Huberman BA, Kaski K, Fortunato S (2015). "Attention decay in science". *Journal of Informetrics*. **9** (4): 734–745. arXiv:1503.01881 (<https://arxiv.org/abs/1503.01881>). Bibcode:2015arXiv150301881D (<https://ui.adsabs.harvard.edu/abs/2015arXiv150301881D>). doi:10.1016/j.joi.2015.07.006 (<https://doi.org/10.1016%2Fj.joi.2015.07.006>). S2CID 10949754 (<https://api.semanticscholar.org/CorpusID:10949754>).
97. Mirowski P (2011). *Science-Mart*. Harvard University Press. pp. 2, 24. ISBN 978-0-674-06113-2.
98. Moeller HG (2006). *Luhmann explained: from souls to systems*. Chicago: Open Court. p. 25. ISBN 0-8126-9598-4. OCLC 68694011 (<https://www.worldcat.org/oclc/68694011>).
99. Luhmann N (1995). *Social systems*. Stanford, CA: Stanford University Press. p. 288. ISBN 978-0-8047-2625-2. OCLC 31710315 (<https://www.worldcat.org/oclc/31710315>).
100. Scheufele DA (September 2014). "Science communication as political communication" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4183176>). *Proceedings of the National Academy of Sciences of the United States of America*. **111** (Supplement 4): 13585–13592. Bibcode:2014PNAS..111S3585S (<https://ui.adsabs.harvard.edu/abs/2014PNAS..111S3585S>). doi:10.1073/pnas.1317516111 (<https://doi.org/10.1073%2Fpnas.1317516111>). PMC 4183176 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4183176>). PMID 25225389 (<https://pubmed.ncbi.nlm.nih.gov/25225389>).
101. Pielke R (2007). *The honest broker : making sense of science in policy and politics*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511818110 (<https://doi.org/10.1017%2FCBO9780511818110>). ISBN 978-0-511-81811-0. OCLC 162145073 (<https://www.worldcat.org/oclc/162145073>).
102. Martin GN, Clarke RM (2017). "Are Psychology Journals Anti-replication? A Snapshot of Editorial Practices" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5387793>). *Frontiers in Psychology*. **8**: 523. doi:10.3389/fpsyg.2017.00523 (<https://doi.org/10.3389%2Ffpsyg.2017.00523>). PMC 5387793 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5387793>). PMID 28443044 (<https://pubmed.ncbi.nlm.nih.gov/28443044>).
103. Yeung AW (2017). "Do Neuroscience Journals Accept Replications? A Survey of Literature" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5611708>). *Frontiers in Human Neuroscience*. **11**: 468. doi:10.3389/fnhum.2017.00468 (<https://doi.org/10.3389%2Ffnhum.2017.00468>). PMC 5611708 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5611708>). PMID 28979201 (<https://pubmed.ncbi.nlm.nih.gov/28979201>).

104. Hubbard R, Vetter DE (1 February 1996). "An empirical comparison of published replication research in accounting, economics, finance, management, and marketing". *Journal of Business Research*. **35** (2): 153–164. doi:10.1016/0148-2963(95)00084-4 (<https://doi.org/10.1016%2F0148-2963%2895%2900084-4>). ISSN 0148-2963 (<https://www.worldcat.org/issn/0148-2963>).
105. "Confirmation bias" (<https://dictionary.apa.org/confirmation-bias>). *APA Dictionary of Psychology*. Washington, DC: American Psychological Association. n.d. Retrieved 2022-02-02.
106. Fanelli D (April 2010). "Do pressures to publish increase scientists' bias? An empirical support from US States Data" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2858206>). *PLOS ONE*. **5** (4): e10271. Bibcode:2010PLoS...510271F (<https://ui.adsabs.harvard.edu/abs/2010PLoS...510271F>). doi:10.1371/journal.pone.0010271 (<https://doi.org/10.1371%2Fjournal.pone.0010271>). PMC 2858206 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2858206>). PMID 20422014 (<https://pubmed.ncbi.nlm.nih.gov/20422014>).
107. Nosek BA, Spies JR, Motyl M (November 2012). "Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10540222>). *Perspectives on Psychological Science*. **7** (6): 615–631. arXiv:1205.4251 (<https://arxiv.org/abs/1205.4251>). doi:10.1177/1745691612459058 (<https://doi.org/10.1177%2F1745691612459058>). PMC 10540222 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10540222>). PMID 26168121 (<https://pubmed.ncbi.nlm.nih.gov/26168121>). S2CID 23602412 (<https://api.semanticscholar.org/CorpusID:23602412>).
108. Everett JA, Earp BD (1 January 2015). "A tragedy of the (academic) commons: interpreting the replication crisis in psychology as a social dilemma for early-career researchers" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4527093>). *Frontiers in Psychology*. **6**: 1152. doi:10.3389/fpsyg.2015.01152 (<https://doi.org/10.3389%2Ffpsyg.2015.01152>). PMC 4527093 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4527093>). PMID 26300832 (<https://pubmed.ncbi.nlm.nih.gov/26300832>).
109. "Research misconduct – The grey area of Questionable Research Practices" (<https://web.archive.org/web/20141031080013/http://www.vib.be/en/news/Pages/Research-misconduct---The-grey-area-of-Questionable-Research-Practices.aspx>). *www.vib.be*. Vlaams Instituut voor Biotechnologie. 30 September 2013. Archived from the original (<http://www.vib.be/en/news/Pages/Research-misconduct---The-grey-area-of-Questionable-Research-Practices.aspx>) on 2014-10-31. Retrieved 2015-11-13.
110. Wicherts JM, Veldkamp CL, Augusteijn HE, Bakker M, van Aert RC, van Assen MA (2016). "Degrees of Freedom in Planning, Running, Analyzing, and Reporting Psychological Studies: A Checklist to Avoid *p*-Hacking" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5122713>). *Frontiers in Psychology*. **7**: 1832. doi:10.3389/fpsyg.2016.01832 (<https://doi.org/10.3389%2Ffpsyg.2016.01832>). PMC 5122713 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5122713>). PMID 27933012 (<https://pubmed.ncbi.nlm.nih.gov/27933012>).
111. Neuroskeptic (November 2012). "The Nine Circles of Scientific Hell" (<https://doi.org/10.1177%2F1745691612459519>). *Perspectives on Psychological Science* (Opinion). **7** (6): 643–644. doi:10.1177/1745691612459519 (<https://doi.org/10.1177%2F1745691612459519>). PMID 26168124 (<https://pubmed.ncbi.nlm.nih.gov/26168124>). S2CID 45328962 (<https://api.semanticscholar.org/CorpusID:45328962>).
112. "Data dredging" (<https://dictionary.apa.org/data-dredging>). *APA Dictionary of Psychology*. Washington, DC: American Psychological Association. n.d. Retrieved 2022-01-09. "The inappropriate practice of searching through large files of information to try to confirm a preconceived hypothesis or belief without an adequate design that controls for possible confounds or alternate hypotheses. Data dredging may involve selecting which parts of a large data set to retain to get specific, desired results."

113. Begley CG (May 2013). "Six red flags for suspect work" (<https://doi.org/10.1038%2F497433a>). *Nature* (Comment article). **497** (7450): 433–434. Bibcode:2013Natur.497..433B (<https://ui.adsabs.harvard.edu/abs/2013Natur.497..433B>). doi:10.1038/497433a (<https://doi.org/10.1038%2F497433a>). PMID 23698428 (<https://pubmed.ncbi.nlm.nih.gov/23698428>). S2CID 4312732 (<https://api.semanticscholar.org/CorpusID:4312732>).
114. Shea C (13 November 2011). "Fraud Scandal Fuels Debate Over Practices of Social Psychology" (<http://chronicle.com/article/As-Dutch-Research-Scandal/129746/>). *The Chronicle of Higher Education*.
115. O'Boyle, Ernest H.; Götz, Martin (2022). "Questionable Research Practices". *Research Integrity: Best Practices for the Social and Behavioral Sciences*. Oxford University Press. pp. 261–294. ISBN 978-0190938550.
116. Glick JL (1992). "Scientific data audit—A key management tool". *Accountability in Research*. **2** (3): 153–168. doi:10.1080/08989629208573811 (<https://doi.org/10.1080%2F08989629208573811>).
117. Fiedler K, Schwarz N (19 October 2015). "Questionable Research Practices Revisited". *Social Psychological and Personality Science*. **7**: 45–52. doi:10.1177/1948550615612150 (<https://doi.org/10.1177%2F1948550615612150>). ISSN 1948-5506 (<https://www.worldcat.org/issn/1948-5506>). S2CID 146717227 (<https://api.semanticscholar.org/CorpusID:146717227>).
118. Fanelli D (May 2009). "How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2685008>). *PLOS ONE*. **4** (5): e5738. Bibcode:2009PLoSO...4.5738F (<https://ui.adsabs.harvard.edu/abs/2009PLoSO...4.5738F>). doi:10.1371/journal.pone.0005738 (<https://doi.org/10.1371%2Fjournal.pone.0005738>). PMC 2685008 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2685008>). PMID 19478950 (<https://pubmed.ncbi.nlm.nih.gov/19478950>).
119. Button, Katherine S.; Ioannidis, John P. A.; Mokrysz, Claire; Nosek, Brian A.; Flint, Jonathan; Robinson, Emma S. J.; Munafò, Marcus R. (May 2013). "Power failure: why small sample size undermines the reliability of neuroscience" (<https://doi.org/10.1038%2Fnrn3475>). *Nature Reviews Neuroscience*. **14** (5): 365–376. doi:10.1038/nrn3475 (<https://doi.org/10.1038%2Fnrn3475>). ISSN 1471-0048 (<https://www.worldcat.org/issn/1471-0048>). PMID 23571845 (<https://pubmed.ncbi.nlm.nih.gov/23571845>). S2CID 455476 (<https://api.semanticscholar.org/CorpusID:455476>).
120. Ioannidis, John P. A.; Stanley, T. D.; Doucouliagos, Hristos (1 October 2017). "The Power of Bias in Economics Research" (<https://doi.org/10.1111%2Fecoj.12461>). *The Economic Journal*. **127** (605): F236–F265. doi:10.1111/ecoj.12461 (<https://doi.org/10.1111%2Fecoj.12461>). ISSN 0013-0133 (<https://www.worldcat.org/issn/0013-0133>). S2CID 158829482 (<https://api.semanticscholar.org/CorpusID:158829482>).
121. Higgins, Julian P. T.; Thompson, Simon G. (15 June 2002). "Quantifying heterogeneity in a meta-analysis" (<https://onlinelibrary.wiley.com/doi/10.1002/sim.1186>). *Statistics in Medicine*. **21** (11): 1539–1558. doi:10.1002/sim.1186 (<https://doi.org/10.1002%2Fsim.1186>). ISSN 0277-6715 (<https://www.worldcat.org/issn/0277-6715>). PMID 12111919 (<https://pubmed.ncbi.nlm.nih.gov/12111919>). S2CID 6319826 (<https://api.semanticscholar.org/CorpusID:6319826>).
122. Moosa IA (2 October 2019). "The fragility of results and bias in empirical research: an exploratory exposition". *Journal of Economic Methodology*. **26** (4): 347–360. doi:10.1080/1350178X.2018.1556798 (<https://doi.org/10.1080%2F1350178X.2018.1556798>). ISSN 1350-178X (<https://www.worldcat.org/issn/1350-178X>). S2CID 158504639 (<https://api.semanticscholar.org/CorpusID:158504639>).
123. Granger CW (1999). *Empirical Modeling in Economics: Specification and Evaluation* (<https://books.google.com/books?id=Lpq49TYm3MMC&pg=PA5>). Cambridge University Press. p. 5. doi:10.1017/CBO9780511492327 (<https://doi.org/10.1017%2FCBO9780511492327>). ISBN 978-0-521-77825-1.

124. Maziarz M (1 December 2021). "Resolving empirical controversies with mechanistic evidence" (<https://doi.org/10.1007%2Fs11229-021-03232-2>). *Synthese*. **199** (3): 9957–9978. doi:10.1007/s11229-021-03232-2 (<https://doi.org/10.1007%2Fs11229-021-03232-2>). ISSN 1573-0964 (<https://www.worldcat.org/issn/1573-0964>). S2CID 236249427 (<https://api.semanticscholar.org/CorpusID:236249427>).
125. Morgan MS, Magnus JR (September 1997). "The experiment in applied econometrics" (<https://onlinelibrary.wiley.com/toc/10991255/1997/12/5>). *Journal of Applied Econometrics*. **12** (5): 459–661. ISSN 1099-1255 (<https://www.worldcat.org/issn/1099-1255>).
126. Van Bavel, Jay J.; Mende-Siedlecki, Peter; Brady, William J.; Reinero, Diego A. (2016). "Contextual sensitivity in scientific reproducibility" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4988618>). *Proceedings of the National Academy of Sciences of the United States of America*. **113** (23): 6454–6459. Bibcode:2016PNAS..113.6454V (<https://ui.adsabs.harvard.edu/abs/2016PNAS..113.6454V>). doi:10.1073/pnas.1521897113 (<https://doi.org/10.1073%2Fpnas.1521897113>). ISSN 0027-8424 (<https://www.worldcat.org/issn/0027-8424>). JSTOR 26470212 (<https://www.jstor.org/stable/26470212>). PMC 4988618 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4988618>). PMID 27217556 (<https://pubmed.ncbi.nlm.nih.gov/27217556>).
127. Bird A (1 December 2021). "Understanding the Replication Crisis as a Base Rate Fallacy" (<https://doi.org/10.1093%2Fbjps%2Faxy051>). *The British Journal for the Philosophy of Science*. **72** (4): 965–993. doi:10.1093/bjps/axy051 (<https://doi.org/10.1093%2Fbjps%2Faxy051>). ISSN 0007-0882 (<https://www.worldcat.org/issn/0007-0882>).
128. Nissen SB, Magidson T, Gross K, Bergstrom CT (December 2016). "Publication bias and the canonization of false facts" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5173326>). *eLife*. **5**: e21451. arXiv:1609.00494 (<https://arxiv.org/abs/1609.00494>). doi:10.7554/eLife.21451 (<https://doi.org/10.7554%2FeLife.21451>). PMC 5173326 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5173326>). PMID 27995896 (<https://pubmed.ncbi.nlm.nih.gov/27995896>).
129. University of California San Diego (May 2021). "A new replication crisis: Research that is less likely to be true is cited more" (<https://phys.org/news/2021-05-replication-crisis-true-cited.html>). *phys.org*. Retrieved 2021-06-14.
130. Serra-Garcia M, Gneezy U (May 2021). "Nonreplicable publications are cited more than replicable ones" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8139580>). *Science Advances*. **7** (21): eabd1705. Bibcode:2021SciA...7.1705S (<https://ui.adsabs.harvard.edu/abs/2021SciA...7.1705S>). doi:10.1126/sciadv.abd1705 (<https://doi.org/10.1126%2Fsciadv.abd1705>). PMC 8139580 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8139580>). PMID 34020944 (<https://pubmed.ncbi.nlm.nih.gov/34020944>).
131. Białek M (January 2018). "Replications can cause distorted belief in scientific progress". *The Behavioral and Brain Sciences*. **41**: e122. doi:10.1017/S0140525X18000584 (<https://doi.org/10.1017%2FS0140525X18000584>). PMID 31064528 (<https://pubmed.ncbi.nlm.nih.gov/31064528>). S2CID 147705650 (<https://api.semanticscholar.org/CorpusID:147705650>).
132. Mede NG, Schäfer MS, Ziegler R, Weißkopf M (January 2021). "The "replication crisis" in the public eye: Germans' awareness and perceptions of the (ir)reproducibility of scientific research" (<https://osf.io/ctpyn/>). *Public Understanding of Science*. **30** (1): 91–102. doi:10.1177/0963662520954370 (<https://doi.org/10.1177%2F0963662520954370>). PMID 32924865 (<https://pubmed.ncbi.nlm.nih.gov/32924865>). S2CID 221723269 (<https://api.semanticscholar.org/CorpusID:221723269>).
133. Letzter R (22 September 2016). "Scientists are furious after a famous psychologist accused her peers of 'methodological terrorism'" (<https://www.businessinsider.com/susan-fiske-methodological-terrorism-2016-9>). *Business Insider*. Retrieved 2020-01-30.
134. "Draft of Observer Column Sparks Strong Social Media Response" (<https://www.psychologicalscience.org/publications/observer/obsonline/draft-of-observer-column-sparks-strong-social-media-response.html>). *APS Observer*. Association for Psychological Science. September 2016. Retrieved 2017-10-04.

135. Fiske ST (31 October 2016). "A Call to Change Science's Culture of Shaming" (<https://www.psychologicalscience.org/observer/a-call-to-change-sciences-culture-of-shaming>). *APS Observer*. **29** (9).
136. Singal J (12 October 2016). "Inside Psychology's 'Methodological Terrorism' Debate" (<http://nymag.com/scienceofus/2016/10/inside-psychologys-methodological-terrorism-debate.html>). *NY Mag*. Retrieved 2017-10-04.
137. Vazire, Simine (2 July 2018). "Implications of the credibility revolution for productivity, creativity, and progress" (<http://journals.sagepub.com/doi/10.1177/1745691617751884>). *Perspectives on Psychological Science*. **13** (4): 411–417. doi:10.1177/1745691617751884 (<https://doi.org/10.1177%2F1745691617751884>). ISSN 1745-6916 (<https://www.worldcat.org/issn/1745-6916>). PMID 29961410 (<https://pubmed.ncbi.nlm.nih.gov/29961410>). S2CID 49647586 (<https://api.semanticscholar.org/CorpusID:49647586>).
138. Korbmacher, Max; Azevedo, Flavio; Pennington, Charlotte R.; et al. (25 July 2023). "The replication crisis has led to positive structural, procedural, and community changes" (<https://doi.org/10.1038%2Fs44271-023-00003-2>). *Communications Psychology*. **1** (1): 1–13. doi:10.1038/s44271-023-00003-2 (<https://doi.org/10.1038%2Fs44271-023-00003-2>). hdl:10852/106350 (<https://hdl.handle.net/10852%2F106350>). ISSN 2731-9121 (<https://www.worldcat.org/issn/2731-9121>).
139. Stroebe W, Strack F (January 2014). "The Alleged Crisis and the Illusion of Exact Replication" (<https://research.rug.nl/en/publications/aa4fb44f-e175-4ca4-b2ea-7ca5a0871c20>). *Perspectives on Psychological Science*. **9** (1): 59–71. doi:10.1177/1745691613514450 (<https://doi.org/10.1177%2F1745691613514450>). PMID 26173241 (<https://pubmed.ncbi.nlm.nih.gov/26173241>). S2CID 31938129 (<https://api.semanticscholar.org/CorpusID:31938129>).
140. Jensen A (7 May 2019). "Replication as Success and Unsuccessful Replication" (<https://cla.umn.edu/philosophy/story/replication-success-and-unsuccessful-replication>). *College of Liberal Arts, Department of Philosophy*. University of Minnesota. Retrieved 2022-05-25.
141. Madhusoodanan J (May 2022). "The overlooked variable in animal studies: why diet makes a difference" (<https://doi.org/10.1038%2Fd41586-022-01393-9>). *Nature*. **605** (7911): 778–779. Bibcode:2022Natur.605..778M (<https://ui.adsabs.harvard.edu/abs/2022Natur.605..778M>). doi:10.1038/d41586-022-01393-9 (<https://doi.org/10.1038%2Fd41586-022-01393-9>). PMID 35606524 (<https://pubmed.ncbi.nlm.nih.gov/35606524>). S2CID 249015202 (<https://api.semanticscholar.org/CorpusID:249015202>).
142. Ioannidis JP (June 2016). "Why Most Clinical Research Is Not Useful" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4915619>). *PLOS Medicine*. **13** (6): e1002049. doi:10.1371/journal.pmed.1002049 (<https://doi.org/10.1371%2Fjournal.pmed.1002049>). PMC 4915619 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4915619>). PMID 27328301 (<https://pubmed.ncbi.nlm.nih.gov/27328301>).
143. Ioannidis JP, Fanelli D, Dunne DD, Goodman SN (October 2015). "Meta-research: Evaluation and Improvement of Research Methods and Practices" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4592065>). *PLOS Biology*. **13** (10): e1002264. doi:10.1371/journal.pbio.1002264 (<https://doi.org/10.1371%2Fjournal.pbio.1002264>). PMC 4592065 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4592065>). PMID 26431313 (<https://pubmed.ncbi.nlm.nih.gov/26431313>).
144. Bach B (8 December 2015). "On communicating science and uncertainty: A podcast with John Ioannidis" (<https://scopeblog.stanford.edu/2015/12/08/on-communicating-science-and-uncertainty-a-podcast-with-john-ioannidis/>). *Scope*. Retrieved 2019-05-20.
145. Gosselin RD (January 2020). "Statistical Analysis Must Improve to Address the Reproducibility Crisis: The ACcess to Transparent Statistics (ACTS) Call to Action". *BioEssays*. **42** (1): e1900189. doi:10.1002/bies.201900189 (<https://doi.org/10.1002%2Fbies.201900189>). PMID 31755115 (<https://pubmed.ncbi.nlm.nih.gov/31755115>). S2CID 208228664 (<https://api.semanticscholar.org/CorpusID:208228664>).

146. Pratt MG, Kaplan S, Whittington R (6 November 2019). "Editorial Essay: The Tumult over Transparency: Decoupling Transparency from Replication in Establishing Trustworthy Qualitative Research" (<https://doi.org/10.1177%2F0001839219887663>). *Administrative Science Quarterly*. **65** (1): 1–19. doi:10.1177/0001839219887663 (<https://doi.org/10.1177%2F0001839219887663>). ISSN 0001-8392 (<https://www.worldcat.org/issn/0001-8392>). S2CID 210537501 (<https://api.semanticscholar.org/CorpusID:210537501>).
147. Aschwanden C (6 December 2018). "Psychology's Replication Crisis Has Made The Field Better" (<https://fivethirtyeight.com/features/psychologys-replication-crisis-has-made-the-field-better/>). *FiveThirtyEight*. Retrieved 2018-12-19.
148. Chartier C, Kline M, McCarthy R, Nuijten M, Dunleavy DJ, Ledgerwood A (December 2018), "The Cooperative Revolution Is Making Psychological Science Better" (<https://www.psychologicalscience.org/observer/the-cooperative-revolution-is-making-psychological-science-better>), *Observer*, **31** (10), retrieved 2018-12-19
149. "Registered Replication Reports" (<http://www.psychologicalscience.org/index.php/replication>). Association for Psychological Science. Retrieved 2015-11-13.
150. Chambers C (20 May 2014). "Psychology's 'registration revolution' " (<https://www.theguardian.com/science/head-quarters/2014/may/20/psychology-registration-revolution>). *The Guardian*. Retrieved 2015-11-13.
151. Lindsay DS (December 2015). "Replication in Psychological Science" (<https://doi.org/10.1177%2F0956797615616374>). *Psychological Science*. **26** (12): 1827–1832. doi:10.1177/0956797615616374 (<https://doi.org/10.1177%2F0956797615616374>). PMID 26553013 (<https://pubmed.ncbi.nlm.nih.gov/26553013>).
152. Benjamin DJ, Berger JO, Johannesson M, Nosek BA, Wagenmakers EJ, Berk R, et al. (January 2018). "Redefine statistical significance" (<https://doi.org/10.1038%2Fs41562-017-0189-z>). *Nature Human Behaviour*. **2** (1): 6–10. doi:10.1038/s41562-017-0189-z (<https://doi.org/10.1038%2Fs41562-017-0189-z>). hdl:10281/184094 (<https://hdl.handle.net/10281%2F184094>). PMID 30980045 (<https://pubmed.ncbi.nlm.nih.gov/30980045>).
153. Lakens D, Adolphi FG, Albers CJ, Anvari F, Apps MA, Argamon SE, et al. (March 2018). "Justify your alpha" (<https://pure.eur.nl/en/publications/a51f8b58-977c-45cc-a977-f1c5839e5afe>). *Nature Human Behaviour*. **2** (3): 168–171. doi:10.1038/s41562-018-0311-x (<https://doi.org/10.1038%2Fs41562-018-0311-x>). hdl:21.11116/0000-0004-9413-F (<https://hdl.handle.net/21.11116%2F0000-0004-9413-F>). ISSN 2397-3374 (<https://www.worldcat.org/issn/2397-3374>). S2CID 3692182 (<https://api.semanticscholar.org/CorpusID:3692182>).
154. Colquhoun D (November 2014). "An investigation of the false discovery rate and the misinterpretation of p-values" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4448847>). *Royal Society Open Science*. **1** (3): 140216. arXiv:1407.5296 (<https://arxiv.org/abs/1407.5296>). Bibcode:2014RSOS....140216C (<https://ui.adsabs.harvard.edu/abs/2014RSOS....140216C>). doi:10.1098/rsos.140216 (<https://doi.org/10.1098%2Frsos.140216>). PMC 4448847 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4448847>). PMID 26064558 (<https://pubmed.ncbi.nlm.nih.gov/26064558>).
155. Colquhoun D (December 2017). "The reproducibility of research and the misinterpretation of p-values" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5750014>). *Royal Society Open Science*. **4** (12): 171085. doi:10.1098/rsos.171085 (<https://doi.org/10.1098%2Frsos.171085>). PMC 5750014 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5750014>). PMID 29308247 (<https://pubmed.ncbi.nlm.nih.gov/29308247>).
156. Colquhoun D (11 October 2016). "The problem with p-values" (<https://aeon.co/essays/it-s-time-for-science-to-abandon-the-term-statistically-significant>). *Aeon Magazine*. Retrieved 2016-12-11.
157. Longstaff C, Colquhoun D. "Calculator for false positive risk (FPR)" (<http://fpr-calc.ucl.ac.uk/>). University College London. version 1.7.

158. Matthews RA (2001). "Why should clinicians care about Bayesian methods?". *Journal of Statistical Planning and Inference*. **94**: 43–58. doi:10.1016/S0378-3758(00)00232-9 (<https://doi.org/10.1016%2FS0378-3758%2800%2900232-9>).
159. Maxwell SE, Lau MY, Howard GS (September 2015). "Is psychology suffering from a replication crisis? What does "failure to replicate" really mean?". *The American Psychologist*. **70** (6): 487–498. doi:10.1037/a0039400 (<https://doi.org/10.1037%2Fa0039400>). PMID 26348332 (<https://pubmed.ncbi.nlm.nih.gov/26348332>).
160. IntHout J, Ioannidis JP, Borm GF, Goeman JJ (August 2015). "Small studies are more heterogeneous than large ones: a meta-meta-analysis" (<https://doi.org/10.1016%2Fj.jclinepi.2015.03.017>). *Journal of Clinical Epidemiology*. **68** (8): 860–869. doi:10.1016/j.jclinepi.2015.03.017 (<https://doi.org/10.1016%2Fj.jclinepi.2015.03.017>). hdl:2066/153978 (<https://hdl.handle.net/2066%2F153978>). PMID 25959635 (<https://pubmed.ncbi.nlm.nih.gov/25959635>).
161. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafò MR (May 2013). "Power failure: why small sample size undermines the reliability of neuroscience" (<https://doi.org/10.1038%2Fnrn3475>). *Nature Reviews. Neuroscience*. **14** (5): 365–376. doi:10.1038/nrn3475 (<https://doi.org/10.1038%2Fnrn3475>). PMID 23571845 (<https://pubmed.ncbi.nlm.nih.gov/23571845>).
162. Greenwald AG (1975). "Consequences of prejudice against the null hypothesis" (https://faculty.washington.edu/agg/pdf/Gwald_PsychBull_1975.OCR.pdf) (PDF). *Psychological Bulletin*. **82** (1): 1–20. doi:10.1037/h0076157 (<https://doi.org/10.1037%2Fh0076157>).
163. Amrhein V, Korner-Nievergelt F, Roth T (2017). "The earth is flat ($p > 0.05$): significance thresholds and the crisis of unreplicable research" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5502092>). *PeerJ*. **5**: e3544. doi:10.7717/peerj.3544 (<https://doi.org/10.7717%2Fpeerj.3544>). PMC 5502092 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5502092>). PMID 28698825 (<https://pubmed.ncbi.nlm.nih.gov/28698825>).
164. "NWO makes 3 million available for Replication Studies pilot" (<https://web.archive.org/web/20160722134546/http://www.nwo.nl/en/news-and-events/news/2016/nwo-makes-3-million-available-for-replication-studies-pilot.html>). *Netherlands Organisation for Scientific Research* (Press release). July 2016. Archived from the original (<http://www.nwo.nl/en/news-and-events/news/2016/nwo-makes-3-million-available-for-replication-studies-pilot.html>) on 2016-07-22.
165. Apple S (22 January 2017). "The Young Billionaire Behind the War on Bad Science" (<https://www.wired.com/2017/01/john-arnold-waging-war-on-bad-science/>). *Wired*.
166. Frank MC, Saxe R (November 2012). "Teaching Replication" (<https://doi.org/10.1177%2F1745691612460686>). *Perspectives on Psychological Science*. **7** (6): 600–604. doi:10.1177/1745691612460686 (<https://doi.org/10.1177%2F1745691612460686>). PMID 26168118 (<https://pubmed.ncbi.nlm.nih.gov/26168118>). S2CID 33661604 (<https://api.semanticscholar.org/CorpusID:33661604>).
167. Grahe JE, Reifman A, Hermann AD, Walker M, Oleson KC, Nario-Redmond M, Wiebe RP (November 2012). "Harnessing the Undiscovered Resource of Student Research Projects" (<https://doi.org/10.1177%2F1745691612459057>). *Perspectives on Psychological Science*. **7** (6): 605–607. doi:10.1177/1745691612459057 (<https://doi.org/10.1177%2F1745691612459057>). PMID 26168119 (<https://pubmed.ncbi.nlm.nih.gov/26168119>).
168. Marwick B, Wang L, Robinson R, Loiselle H (22 October 2019). "How to Use Replication Assignments for Teaching Integrity in Empirical Archaeology" (<https://doi.org/10.1017%2Faap.2019.38>). *Advances in Archaeological Practice*. **8**: 78–86. doi:10.1017/aap.2019.38 (<https://doi.org/10.1017%2Faap.2019.38>).

169. Everett JA, Earp BD (1 January 2015). "A tragedy of the (academic) commons: interpreting the replication crisis in psychology as a social dilemma for early-career researchers" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4527093>). *Frontiers in Psychology*. **6**: 1152. doi:10.3389/fpsyg.2015.01152 (<https://doi.org/10.3389%2Fpsyg.2015.01152>). PMC 4527093 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4527093>). PMID 26300832 (<https://pubmed.ncbi.nlm.nih.gov/26300832>).
170. Quintana DS (September 2021). "Replication studies for undergraduate theses to improve science and education". *Nature Human Behaviour* (World View article). **5** (9): 1117–1118. doi:10.1038/s41562-021-01192-8 (<https://doi.org/10.1038%2Fs41562-021-01192-8>). PMID 34493847 (<https://pubmed.ncbi.nlm.nih.gov/34493847>). S2CID 237439956 (<https://api.semanticscholar.org/CorpusID:237439956>).
171. University of Cambridge (April 2022). "'Robot scientist' Eve finds that less than one-third of scientific results are reproducible" (<https://techxplore.com/news/2022-04-robot-scientist-eve-one-third-scientific.html>). *Techxplore*. Retrieved 2022-05-15.
172. Roper K, Abdel-Rehim A, Hubbard S, Carpenter M, Rzhetsky A, Soldatova L, King RD (April 2022). "Testing the reproducibility and robustness of the cancer biology literature by robot" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8984295>). *Journal of the Royal Society, Interface*. **19** (189): 20210821. doi:10.1098/rsif.2021.0821 (<https://doi.org/10.1098%2Frif.2021.0821>). PMC 8984295 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8984295>). PMID 35382578 (<https://pubmed.ncbi.nlm.nih.gov/35382578>).
173. Chambers C (10 June 2014). "Physics envy: Do 'hard' sciences hold the solution to the replication crisis in psychology?" (<https://www.theguardian.com/science/head-quarters/2014/jun/10/physics-envy-do-hard-sciences-hold-the-solution-to-the-replication-crisis-in-psychology>). *The Guardian*.
174. Kahneman D (2014). "A New Etiquette for Replication". *Social Psychology* (Commentary). Commentaries and Rejoinder on. **45** (4): 310–311. doi:10.1027/1864-9335/a000202 (<https://doi.org/10.1027%2F1864-9335%2Fa000202>).
175. Makel MC, Plucker JA, Hegarty B (November 2012). "Replications in Psychology Research: How Often Do They Really Occur?" (<https://doi.org/10.1177%2F1745691612460688>). *Perspectives on Psychological Science*. **7** (6): 537–542. doi:10.1177/1745691612460688 (<https://doi.org/10.1177%2F1745691612460688>). PMID 26168110 (<https://pubmed.ncbi.nlm.nih.gov/26168110>).
176. Uhlmann, Eric Luis; Ebersole, Charles R.; Chartier, Christopher R.; Errington, Timothy M.; Kidwell, Mallory C.; Lai, Calvin K.; McCarthy, Randy J.; Riegelman, Amy; Silberzahn, Raphael; Nosek, Brian A. (September 2019). "Scientific Utopia III: Crowdsourcing Science" (<https://doi.org/10.1177%2F1745691619850561>). *Perspectives on Psychological Science*. **14** (5): 711–733. doi:10.1177/1745691619850561 (<https://doi.org/10.1177%2F1745691619850561>). ISSN 1745-6916 (<https://www.worldcat.org/issn/1745-6916>). PMID 31260639 (<https://pubmed.ncbi.nlm.nih.gov/31260639>).
177. Forscher, Patrick S.; Wagenmakers, Eric-Jan; Coles, Nicholas A.; Silan, Miguel Alejandro; Dutra, Natália; Basnight-Brown, Dana; IJzerman, Hans (May 2023). "The Benefits, Barriers, and Risks of Big-Team Science" (<http://journals.sagepub.com/doi/10.1177/17456916221082970>). *Perspectives on Psychological Science*. **18** (3): 607–623. doi:10.1177/17456916221082970 (<https://doi.org/10.1177%2F17456916221082970>). ISSN 1745-6916 (<https://www.worldcat.org/issn/1745-6916>). PMID 36190899 (<https://pubmed.ncbi.nlm.nih.gov/36190899>). S2CID 236816530 (<https://api.semanticscholar.org/CorpusID:236816530>).
178. Munafò MR, Davey Smith G (January 2018). "Robust research needs many lines of evidence" (<https://doi.org/10.1038%2Fd41586-018-01023-3>). *Nature*. **553** (7689): 399–401. Bibcode:2018Natur.553..399M (<https://ui.adsabs.harvard.edu/abs/2018Natur.553..399M>). doi:10.1038/d41586-018-01023-3 (<https://doi.org/10.1038%2Fd41586-018-01023-3>). PMID 29368721 (<https://pubmed.ncbi.nlm.nih.gov/29368721>).

179. Wallot S, Kelty-Stephen DG (1 June 2018). "Interaction-Dominant Causation in Mind and Brain, and Its Implication for Questions of Generalization and Replication" (<https://doi.org/10.1007%2Fs11023-017-9455-0>). *Minds and Machines*. **28** (2): 353–374. doi:10.1007/s11023-017-9455-0 (<https://doi.org/10.1007%2Fs11023-017-9455-0>). hdl:21.11116/0000-0001-AC9C-E (<https://hdl.handle.net/21.11116%2F0000-0001-AC9C-E>). ISSN 1572-8641 (<https://www.worldcat.org/issn/1572-8641>).
180. Tierney W, Hardy JH, Ebersole CR, Leavitt K, Viganola D, Clemente EG, et al. (1 November 2020). "Creative destruction in science" (<https://doi.org/10.1016%2Fj.obhdp.2020.07.002>). *Organizational Behavior and Human Decision Processes*. **161**: 291–309. doi:10.1016/j.obhdp.2020.07.002 (<https://doi.org/10.1016%2Fj.obhdp.2020.07.002>). hdl:2066/228242 (<https://hdl.handle.net/2066%2F228242>). ISSN 0749-5978 (<https://www.worldcat.org/issn/0749-5978>). S2CID 224979451 (<https://api.semanticscholar.org/CorpusID:224979451>).
181. Tierney W, Hardy J, Ebersole CR, Viganola D, Clemente EG, Gordon M, et al. (1 March 2021). "A creative destruction approach to replication: Implicit work and sex morality across cultures" (<https://doi.org/10.1016%2Fj.jesp.2020.104060>). *Journal of Experimental Social Psychology*. **93**: 104060. doi:10.1016/j.jesp.2020.104060 (<https://doi.org/10.1016%2Fj.jesp.2020.104060>). hdl:10037/24275 (<https://hdl.handle.net/10037%2F24275>). ISSN 0022-1031 (<https://www.worldcat.org/issn/0022-1031>). S2CID 229028797 (<https://api.semanticscholar.org/CorpusID:229028797>).
182. Delios A, Clemente EG, Wu T, Tan H, Wang Y, Gordon M, et al. (July 2022). "Examining the generalizability of research findings from archival data" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9335312>). *Proceedings of the National Academy of Sciences of the United States of America*. **119** (30): e2120377119. Bibcode:2022PNAS..11920377D (<https://ui.adsabs.harvard.edu/abs/2022PNAS..11920377D>). doi:10.1073/pnas.2120377119 (<https://doi.org/10.1073%2Fpnas.2120377119>). PMC 9335312 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9335312>). PMID 35858443 (<https://pubmed.ncbi.nlm.nih.gov/35858443>).
183. Ince DC, Hatton L, Graham-Cumming J (February 2012). "The case for open computer programs" (<https://doi.org/10.1038%2Fnature10836>). *Nature*. **482** (7386): 485–488. Bibcode:2012Natur.482..485I (<https://ui.adsabs.harvard.edu/abs/2012Natur.482..485I>). doi:10.1038/nature10836 (<https://doi.org/10.1038%2Fnature10836>). PMID 22358837 (<https://pubmed.ncbi.nlm.nih.gov/22358837>).
184. Vuong QH (January 2018). "The (ir)rational consideration of the cost of science in transition economies" (<https://doi.org/10.1038%2Fs41562-017-0281-4>). *Nature Human Behaviour*. **2** (1): 5. doi:10.1038/s41562-017-0281-4 (<https://doi.org/10.1038%2Fs41562-017-0281-4>). PMID 30980055 (<https://pubmed.ncbi.nlm.nih.gov/30980055>). S2CID 46878093 (<https://api.semanticscholar.org/CorpusID:46878093>).
185. Junk TR, Lyons L (21 December 2020). "Reproducibility and Replication of Experimental Particle Physics Results". *Harvard Data Science Review*. **2** (4). arXiv:2009.06864 (<https://arxiv.org/abs/2009.06864>). doi:10.1162/99608f92.250f995b (<https://doi.org/10.1162%2F99608f92.250f995b>). S2CID 221703733 (<https://api.semanticscholar.org/CorpusID:221703733>).
186. Ioannidis JP (February 2016). "Anticipating consequences of sharing raw data and code and of awarding badges for sharing". *Journal of Clinical Epidemiology* (Commentary). **70**: 258–260. doi:10.1016/j.jclinepi.2015.04.015 (<https://doi.org/10.1016%2Fj.jclinepi.2015.04.015>). PMID 26163123 (<https://pubmed.ncbi.nlm.nih.gov/26163123>).

Further reading

- Bastian H (5 December 2016). "Reproducibility Crisis Timeline: Milestones in Tackling Research Reliability" (<https://blogs.plos.org/absolutely-maybe/2016/12/05/reproducibility-crisis-timeline-milestones-in-tackling-research-reliability/>). *Absolutely Maybe*. Retrieved 2019-06-05.