Editorial

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Practicing Good Laboratory Hygiene, Even in a Pandemic

Psychological Science
2020, Vol. 31(5) 483–487

DOI: 10.1177/0956797620920547
www.psychologicalscience.org/PS
SAGE

To say that our consciousness about hygiene has been raised in recent weeks and months is to state the obvious. Who would have thought that early 2020 would see full-length instructional videos on handwashing techniques, set to catchy tunes? More seriously, who would have thought that whole societies would be placed under lockdown orders? Yet these are the new realities of life in the time of coronavirus disease 2019 (COVID-19; with apologies to Gabriel García Márquez, 1985/1988). The implications of the COVID-19 pandemic have touched every aspect of life—and psychological research is no exception. With entire universities moving to remote instruction and virtually all other functions occurring online, researchers are facing an unexpected and sudden end to on-site, in-person data collection. For psychological scientists, this moment brings both promise and peril.

The COVID-19 pandemic offers unique opportunities to advance psychological science by, for example, illuminating the impact of global stressors on human behavior (e.g., Talarico & Rubin, 2003). Further, as scientists who study the human condition, psychologists are uniquely qualified to offer empirically grounded advice about how events such as the COVID-19 pandemic can impact people socially, emotionally, and cognitively. To facilitate dissemination of this work, *Psychological Science* is expediting the review of COVID-19-related submissions and fast-tracking publications of accepted manuscripts.

The pandemic also brings the peril of placing research expedience over rigor and transparency. In our rush to understand the impacts of the pandemic, we may be tempted to relax our vigilance against poor research practices. Moreover, while the pandemic slows or even halts many research projects, the real and perceived time pressures inherent in the lives of academic researchers will not abate. Honors, master's, and doctoral theses must be defended; job applications must be compiled; funding proposals and progress reports must be submitted; and reappointment, tenure, and promotion packages must be readied for reviewers. All of these activities require that data be collected, and a major measure of the success of our endeavors is the number of articles published. With no promise of resumption of on-site, in-person data collection on the horizon for many researchers, and the potential for long-term restrictions as we await a vaccine, the temptation to rescue incomplete or marginally significant data sets is almost certain to arise.

Understanding the Peril

The unfortunate reality is that many of the practices that psychological researchers may undertake in light of the unexpected hiatus of their research programs are likely to negatively impact the integrity of the science. The concern is that many of these strategies directly inflate the false-discovery rate—the likelihood of obtaining a statistically significant outcome when no real relation exists (i.e., when the null hypothesis should not be rejected). Any time a decision is made that alters the intended design or analysis of a study, there is a risk of introducing a circularity (e.g., Kriegeskorte, Simmons, Bellgowan, & Baker, 2009), wherein the significance, and in extreme cases the effect size itself, is exaggerated. Even a practice as seemingly innocuous as checking both means and medians during analysis and then reporting the "better" measure inflates the false-discovery rate. This and other seemingly equally innocuous research practices have been commonplace historically. Their impacts stand out in stark relief in failures to replicate large sections of the literature (Open Science Collaboration, 2015).

As an illustration of the potential problem, consider a hypothetical project that was partially through the intended data collection before being halted by the COVID-19 crisis but that was far enough along to have sufficient data for analysis. This situation engenders a difficult and problematic decision for the researchers: If the analyses come back significant, will they stop the project shy of the intended number of observations and proceed to publication of the results? If the results are in the right direction but not yet significant, will they wait to add more data? Although either option seems logical, under the circumstances, the researchers have actually engaged in *optional stopping*—using the outcome of an analysis to decide whether more data are needed (e.g., Simmons, Nelson, & Simonsohn, 2011). Optional stopping inflates false-discovery rates by rigging the results, akin to gambling on a coin flip with the rule that if the coin comes up tails, I win, and if it comes up heads, we flip again.

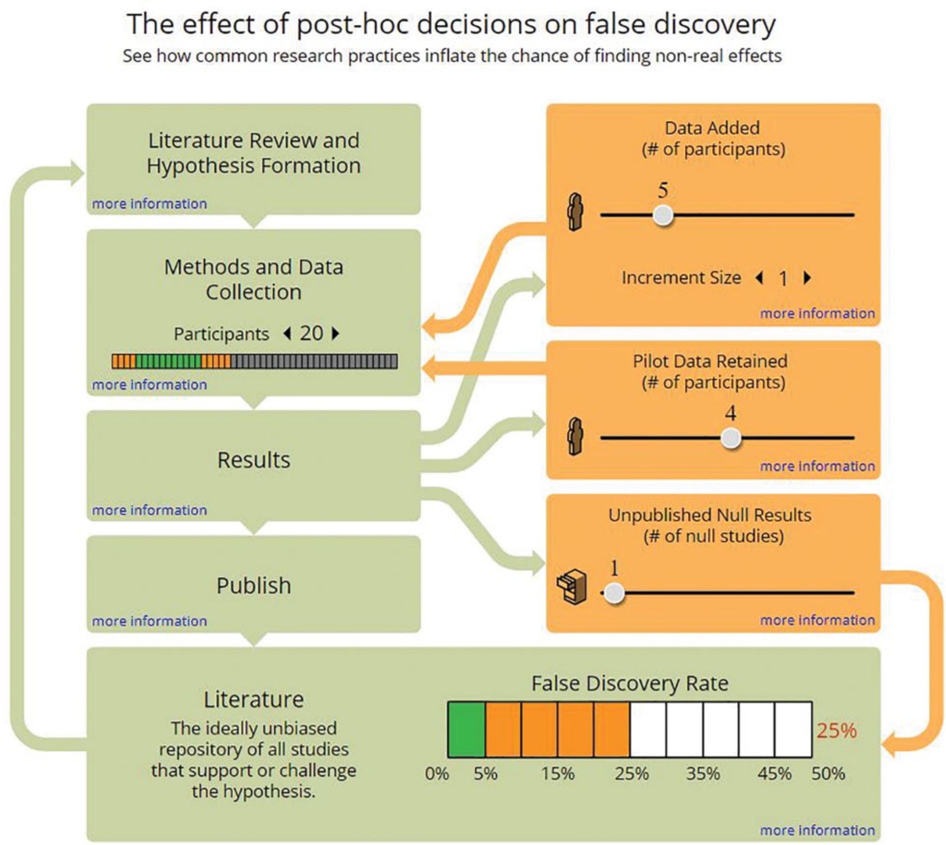

Fig. 1. Screenshot of the interactive applet (see www.bigcogsci.com/post-hoc.html) that can be used to determine the effect of post hoc decisions on false discovery.

Now consider another hypothetical project that is just a few participants short of the intended sample size. In this situation, researchers might consider including data from the final one or two (or even more) individuals previously designated as pilot participants. Again, the practice seems reasonable because the participants were drawn from the same population and were tested under the same design. However, it is highly likely that these were the final pilot participants for a reason—because their data showed at least a hint of the hypothesized effect (akin to the adage that you always find something in the last place you look for it). Yet the likely result of this practice is that the ultimate estimate of the effect will be inflated in both size and statistical significance. If you will, this is like going fishing but starting with a few fish already in your bucket—a head start that will give a false impression of a good day on the lake.

Concrete Illustrations From *Airport Scanner*

The impact of "logical" and "reasonable" practices such as those just described can be illustrated with a recent project using a massive set of real human performance data. The project permits quantification of the impact of these and other practices on the false-discovery rate. In this project, Kravitz and Mitroff (2020) analyzed "big data" from the mobile game *Airport Scanner* (https://www.airportscannergame.com; Kedlin Company, 2018). In this game, players take on the role of an aviation security screener and search through simulated bags to find prohibited items. The game was extremely popular, generating a data set of over 3.7 billion trials from over 15 million mobile devices. The data have been used to address issues in psychological research (e.g., Ericson, Kravitz, & Mitroff, 2017; Mitroff et al., 2015; Mitroff & Sharpe, 2017). Kravitz and Mitroff analyzed thousands of independent samples from the game and used them to demonstrate that even a subset of these seemingly harmless practices, alone or in combination, can inflate the false-discovery rate. Through their analysis, it is easy to see how the psychological literature could find itself with a replication crisis. To help illustrate the scope of the problem and how the practices interact, see Figure 1 (an interactive version of the applet shown in the figure is available at www.bigcogsci.com/post-hoc.html). This applet allows users to explore the impact of some common research practices on inflation of the false-discovery rate. Critically, the project does more than illustrate the negative impact of a number of questionable research practices: It also provides guidance on the size of the adjustment in statistical analyses needed to correct for the inflation of the false-discovery rate (Kravitz & Mitroff, 2020).

Conclusions

In times of great uncertainty, it is not always obvious how to balance the demands of discovery and research productivity while maintaining rigor in our research practices (e.g., maintaining good laboratory hygiene). In general, the right path is to simply avoid questionable practices altogether. We may view this as the equivalent of practicing social distancing. However, in some situations—such as those arising from the COVID-19 pandemic—questionable research practices may seem unavoidable. In such cases, it is imperative that we be fully transparent in reporting the research decisions we made and why we made them. Moreover, aided by tools such as the applet based on the massive *Airport Scanner* data set (Kravitz & Mitroff, 2020), it is incumbent on us to adjust our statistical analyses to remove the inflation of the false-discovery rate. In short, if we find it necessary to stray from our original research designs in ways that add a circularity to the analyses, then "$p < .05$" is no longer a justified threshold for significance.

The entire world is hopeful that the COVID-19 pandemic will soon be a distant memory. Yet we are living it at the moment. In terms of our research, the decisions we make now will impact the literature for years, even decades, to come. We must be vigilant lest expediency stain the future. Fortunately, we have a growing body of materials that illustrate the need to maintain good laboratory hygiene (i.e., the impact of questionable research practices) and that provide guidance on how to achieve it (i.e., suggestions for improving research practices). The *Airport Scanner* applet introduced in this Editorial is a source of both. In the Appendix are links to other resources previously published by the Association for Psychological Science, the Center for Open Science, and others that we can use to improve the quality of psychological science now and in the future. As long as we are appropriately reflective—and transparent—we can maintain positive momentum, even as we shelter in place.

Dwight J. Kravitz
Department of Psychological and Brain Sciences, The George Washington University
Stephen R. Mitroff
Department of Psychological and Brain Sciences, The George Washington University
Patricia J. Bauer
Department of Psychology, Emory University

Appendix

There has been a recent explosion in the study of and resources for reproducible and transparent research. Here is a nonexhaustive list of some of these resources that are particularly relevant to psychological research. The list is organized by the phase of a research project at which the resource or consideration should be applied.

Design

During the initial design of a project, there are three important issues to consider before data collection has begun: (a) the ethics of the design, (b) the amount of data needed to test for the hypothesized results (power), and (c) preregistering the design to help avoid questionable research practices such as those discussed above.

Ethical design.

- The World Medical Association Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects (2020; https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/)

Power analysis.

- "Researchers' Intuitions About Power in Psychological Research" (2016) by Marjan Bakker, Chris H. J. Hartgerink, Jelte M. Wicherts, and Han L. J. van der Maas (https://doi.org/10.1177/0956797616647519)
- "The Statistical Power of Abnormal-Social Psychological Research: A Review" (1962) by Jacob Cohen (https://doi.org/10.1037/h0045186)
- "Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors" (2014) by Andrew Gelman and John Carlin (https://doi.org/10.1177/1745691614551642)

Preregistration.

- "Research Preregistration 101" (2016) by D. Stephen Lindsay, Daniel J. Simons, and Scott O. Lilienfeld (https://www.psychologicalscience.org/observer/research-preregistration-101)
- "Ensuring the Quality and Specificity of Preregistrations" (2020) by Marjan Bakker, Coosje L. S. Veldkamp, Marcel A. L. M. van Assen, Elise A. V. Crompvoets, How Hwee Ong, Brian A. Nosek, Courtney K. Soderberg, David Mellor, and Jelte, M. Wicherts (https://doi.org/10.31234/osf.io/cdgyh)
- "OSF Preregistration" (2020; https://osf.io/prereg/)

Data analysis

Once data sets have been collected, it is important to apply the appropriate statistical analyses. Below are a list of resources that detail modern general statistical approaches and the New Statistics approach advocated by the Association for Psychological Science (APS). Critically, no matter the analyses used, they cannot be allowed to cause alterations to the design of the project (e.g., number of participants run, inclusion or exclusion criteria), or false-discovery rates will be inflated.

General approaches and resources.

- "JASP: A Fresh Way to Do Statistics" (2019) by the JASP Team (https://jasp-stats.org/)
- "A Tutorial on a Practical Bayesian Alternative to Null-Hypothesis Significance Testing" (2011) by Michael E. J. Masson (https://doi.org/10.3758/s13428-010-0049-5)
- "Making 'Null Effects' Informative: Statistical Techniques and Inferential Frameworks" (2018) by Christopher Harms and Daniël Lakens (https://psyarxiv.com/48zca)

New statistics.

- *Introduction to the New Statistics: Estimation, Open Science, & Beyond* (2017) by Geoff Cumming and Robert Calin-Jageman (https://thenewstatistics.com/itns/)
- "The New Statistics: Why and How" (2014) by Geoff Cumming (https://doi.org/10.1177/0956797613504966)
- *The New Statistics: Estimation and Research Integrity* (2014) by Geoff Cumming (https://www.psychologicalscience.org/members/new-statistics)

Post hoc inflation of false-discovery rate.

- "Quantifying, and Correcting For, the Impact of Questionable Research Practices on False Discovery Rates in Psychological Science" (2020) by Dwight J. Kravitz and Stephen R. Mitroff (https://doi.org/10.31234/osf.io/fu9gy)

Manuscript preparation

Preparing a manuscript for submission to a journal entails almost as many decisions as designing the study and analyzing the data. These resources provide guidance on what information should be included in the report.

- "Constraints on Generality (COG): A Proposed Addition to All Empirical Papers" (2017) by Daniel J. Simons, Yuichi Shoda, and D. Stephen Lindsay (https://doi.org/10.1177/1745691617708630)
- "Preparing Manuscripts for Journal Publication in Psychology Articles: A Guide for New Authors" (2010) by the American Psychological Association (https://www.apa.org/pubs/authors/new-author-guide.pdf)
- "Manuscript Structure, Style, and Content Guidelines" (2020) by the Association for Psychological Science (https://www.psychologicalscience.org/publications/ms-structure-guidelines)

Manuscript submission

Once a project reaches the point of submission for publication, it is important to make sure several best practices are followed: The accuracy of the reported statistics should be checked, methods and analyses should be reported as transparently as possible, and, whenever possible, data should be made publicly available.

Statistics checking.

- "statcheck: Extract Statistics From Articles and Recompute p-Values" (web application; 2016) by S. C. Rife, M. B. Nuijten, & S. Epskamp (http://statcheck.io/)

Transparent reporting.

- "Business Not as Usual" (2014) by Eric Eich (https://doi.org/10.1177/0956797613512465)
- "Sharing Data and Materials in *Psychological Science*" (2017) by D. Stephen Lindsay (https://doi.org/10.1177/0956797617704015)
- "Transparent Science: A More Credible, Reproducible, and Publishable Way to Do Science" (2018) by David Mellor, Simine Vazire, and D. Stephen Lindsay (https://doi.org/10.31234/osf.io/7wkdn)

Making data publicly available.

- "Replication in Psychological Science" (2015) by D. Stephen Lindsay (https://doi.org/10.1177/0956797615616374)
- "Sharing Data and Materials in *Psychological Science*" (2017) by D. Stephen Lindsay (https://doi.org/10.1177/0956797617704015)
- The Open Science Framework (osf.io)

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

References

References marked with an asterisk indicate articles included in the Appendix.

*American Psychological Association. (2010). *Preparing manuscripts for journal publication in psychology articles: A guide for new authors*. Retrieved from https://www.apa.org/pubs/authors/new-author-guide.pdf

*Association for Psychological Science. (2020). *Manuscript structure, style, and content guidelines*. Retrieved from https://www.psychologicalscience.org/publications/ms-structure-guidelines

*Bakker, M., Veldkamp, C. L. S., van Assen, M. A. L. M., Crompvoets, E. A. V., Ong, H. H., Nosek, B. A, . . . Wicherts, J. M. (2020). Ensuring the quality and specificity of preregistrations. *PsyArXiv Preprints*. doi:10.31234/osf.io/cdgyh

*Bakker, M., Hartgerink, C. H. J., Wicherts, J. M., & van der Mass, H. L. J. (2016). Researchers' intuitions about power in psychological research. *Psychological Science, 27*, 1069–1077. doi:10.1177/0956797616647519

*Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal Social Psychology, 65*, 145-153. (https://doi.org/10.1037/h0045186)

*Cumming, G. (2014). *The new statistics: Estimation and research integrity*. Retrieved from https://www.psychologicalscience.org/members/new-statistics

*Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*, 7–29. doi:10.1177/0956797613504966

*Cumming, G., & Calin-Jageman, R. (2017). *Introduction to the new statistics: Estimation, open science, & beyond*. Abingdon, England: Routledge.

*Eich, E. (2014). Business not as usual. *Psychological Science, 25,* 3–6. doi:10.1177/0956797613512465

Ericson, J. M., Kravitz, D. J., & Mitroff, S. R. (2017). Visual search: You are who you are (+ A Learning Curve). *Perception, 46*, 1434–1441. doi:10.1177/0301006617721091

García Márquez, G. (1988). *Love in the time of cholera* (E. Grossman, Trans.). New York, NY: Knopf. (Original work published 1985)

*Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing Type S (Sign) and Type M (Magnitude) errors. *Perspectives on Psychological Science, 9*, 641–651. doi:10.1177/1745691614551642

*Harms, C., & Lakens, D. (2018). Making 'null effects' informative: Statistical techniques and inferential frameworks. *PsyArXiv Preprints*. Retrieved from https://psyarxiv.com/48zca

*JASP Team. (2019). *JASP—A fresh way to do statistics*. Retrieved from https://jasp-stats.org/

Kedlin Company. (2018). Airport Scanner [Mobile application software]. Retrieved from http://itunes.apple.com

*Kravitz, D. J., & Mitroff, S. R. (2020). Quantifying, and correcting for, the impact of questionable research practices on false discovery rates in psychological science. *PsyArXiv Preprints*. doi:10.31234/osf.io/fu9gy

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience, 12*, 535–540. doi:10.1038/nn.2303

*Lindsay, D. S. (2015). Replication in Psychological Science. *Psychological Science, 26,* 1827–1832. doi:10.1177/0956797615616374

*Lindsay, D. S. (2017). Sharing data and materials in *Psychological Science*. *Psychological Science, 28*, 699–702. doi:10.1177/0956797617704015

*Lindsay, D. S., Simons, D. J., & Lilienfeld, S. O. (2016, December). Research preregistration 101. *Observer*. Retrieved from https://www.psychologicalscience.org/observer/research-preregistration-101

*Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43*, 679–690. doi:10.3758/s13428-010-0049-5

*Mellor, D., Vazire, S., & Lindsay, D. S. (2018). Transparent science: A more credible, reproducible, and publishable way to do science. *PsyArXiv Preprints*. doi:10.31234/osf.io/7wkdn

Mitroff, S. R., Biggs, A. T., Adamo, S. H., Dowd, E. W., Winkle, J., & Clark, K. (2015). What can 1 billion trials tell us about visual search? *Journal of Experimental Psychology: Human Perception and Performance, 41,* 1–5. doi:10.1037/xhp0000012

Mitroff, S. R., & Sharpe, B. (2017). Using big data to solve real problems through academic and industry partnerships. *Current Opinion in Behavioral Sciences, 18*, 91–96. doi:10.1016/j.cobeha.2017.09.013

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), Article aac4716. doi:10.1126/science.aac4716

*The Open Science Framework. (2020). *OSF Preregistration*. Retrieved from osf.io/prereg/

*The Open Science Framework. (2020). *OSFHOME*. Retrieved from osf.io

*Rife, S. C., Nuijten, M. B., & Epskamp, S. (2016). statcheck: Extract statistics from articles and recompute p-values [web application]. Retrieved from http://statcheck.io

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

*Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science, 12,* 1123–1128. doi:10.1177/1745691617708630

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science*, *14,* 455–461. doi:10.1111/1467-9280.02453

*The World Medical Association. (2020). *WMA Declaration of Helsinki – ethical principles for medical research involving human subjects*. Retrieved from https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/