Editorial

Aaron S. Benjamin
University of Illinois

The first volume of the *Journal of Experimental Psychology*, published in 1916 under the editorship of John Watson, published more articles than all of the other American Psychological Association journals combined. It featured articles on perception *A preliminary study of tonal volume*, attention (*Outline as a condition of attention*), memory (*The factors affecting a permanent impression developed through repetition*), action (*The constant error of touch localization*), and problem solving (*A graded series of geometrical puzzles*). It also included articles on testing (*Report of psychological tests at Reed College*), and methodological advances (*A new method of heterochromatic photometry*). As the discipline of psychology grew, the American Psychological Association divided the journal, and expanded their stable of journals to more than 100—of which the current *Journal of Experimental Psychology: Learning, Memory, and Cognition* (JEP:LMC) is the largest, having published 413 articles in 2017. Yet as cognitive psychology and its allied disciplines have grown, it is notable that most of those articles from 1916 would be a comfortable fit, in substance if not style, for the journal of 2019.

It is gratifying to see that our discipline is mature enough to have subject matter that has remained important over a century of investigation. But one cannot help but wonder whether that persistence also reflects a genuine lack of progress. Some have worried that a lack of grand theoretical sophistication has turned cognitive psychology into an investigation of *phenomena*, with only fractal progress, never approaching a unified understanding of complex cognitive phenomena (Newell, 1973) or important real-world problems (Neisser, 1976). Others have argued that cognitive psychology should be more, not less, informed by specific applied problems and avoid a fetishization of theory (Bartlett, 1932; Broadbent, 1980). Recent concerns are more prosaic: Maybe much of what we are publishing is just factually wrong (Pashler & Wagenmakers, 2012; Simmons, Nelson, & Simonsohn, 2011)?

I will admit to sympathy for the perspective that suboptimal and inefficient methodological behaviors are holding us back from producing the most advanced psychological science. The combined *Journals of Experimental Psychology* published more than 1,000 papers in 2017, with only narrow and mostly unrewarded quality control and oversight. Domain expertise has narrowed as it has deepened, meaning that research areas may not be cross-pollinating effectively. As subfields multiply and cleave off, some have undoubtedly settled into persisting nonoptimal patterns. Ready availability of computer software for running experiments and for analyzing complex designs, not to mention access to vast numbers of online subjects, reduce the barrier to entry and make complex designs more feasible to administer than ever. And competition is stiff to publish in premier journals like this one. Undoubtedly this combination of circumstances, along with incentives to publish eye-catching research—and a lot of it—should make us wary.

Editorial Policy and Emphasis

So I would like to nudge contributors to JEP:LMC in the direction of producing more replicable and reproducible research. This will mean little to no change in practice for many researchers. For others, it will mean paying closer attention to, and being more transparent about, issues of design and analysis. For the Journal, it means we will be promoting good scientific hygiene by holding individual papers to a high methodological standard. JEP:LMC will encourage within-paper replications, especially for exploratory findings. We will be open to publishing **replications** of important work, although it is important to understand that the standards for publishing replications will be different from publishing an original piece of research. The key, as always, is the information value of the findings. Information value is a joint function of the surprisingness of the finding and its validity—a replication must be methodologically and analytically sound and must advance our understanding by refuting a commonly accepted result or by clarifying a questionable one. **Null results**, in replications and throughout the Journal, are acceptable when you can convince the editorial team of the importance and conclusiveness of the result as well as the broad acceptability of the method as a fair test.

I would like the journal to provide more tools but few proscriptions for researchers to improve rigor. I encourage authors to make **data, materials, and analysis scripts openly available** when it is feasible to do so. (And I will also use this moment to remind the reader that it is APA policy to make your data available upon request, even if it is not publicly available.) I encourage **preregistration**, and am opening the Journal to **Registered Reports** (more on this below). Icons indicating these practices will be given, although I have a slight distaste for the callow term of art badges and prefer to emphasize their value as direct links that make supporting material easily accessible from the online article.

Estimates of power and their implications for sample size are a critical part of the endeavor to improve the replicability and validity of our science. **Sample sizes** and number of repeated measures should be straightforwardly justifiable, even in cases in which logistics, resources, or psychological limitations of the subjects render high power difficult to achieve. It is preferable to have transparently moderate power to detect an effect of realistic

Correspondence concerning this article should be addressed to Aaron S. Benjamin, Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, IL 61820. E-mail: asbenjam@illinois.edu

size than exceptional power to detect an optimistic but unrealistic effect.

Different approaches are acceptable for convincing readers of the replicability of your work. Many of the strategies listed above are intended to reduce the opportunities to avail oneself of the garden of forking paths (Gelman & Loken, 2013; Simmons et al., 2011) when making decisions about how to analyze and report data. A less discussed but extremely effective strategy is simply to run experiments with **simple, uncluttered designs**. It is always worth asking whether the benefit of including an additional variable outweighs the costs of having the freedom to make decisions about the treatment of that variable. When that decision is made on the fly, after data are collected, it subjects one to concerns about capitalizing upon sampling error. When it is made ahead of time—for example, in a preregistration—it may be made naively. Simple designs reduce these choice points. Design specifications also touch upon the issue of interactions, about which I have more to say below.

A related point can be made about outlier policy and analytic decisions more generally. Subjects and data points should be omitted from analysis only when that decision can be justified, and the biasing effects of omission must be recognized. Manipulation checks that reveal a failure to adhere to basic preconditions of the task (staying awake, being sober) or to master the instructions (reversed responding, gazing at the wrong portion of the screen) are reasonable, but poor performance or slow responding might not be in the absence of additional evidence. **Statistical tools that are robust to outliers** (such as nonparametric approaches and cross-validation) are preferred over adjustments to the data set to contort the data into one's preferred analytic tool. In general, it is important to be aware, and to be explicit, about the distinction between tests that are confirmatory and those that are exploratory.

Interactions

Interactions are meaningful only under limited conditions (Loftus, 1978). If you write an article in which you want to claim that a manipulation affects *this* more than *that*, you should be prepared to defend the validity of that inference. There are many techniques for doing so, and I will not presume to dictate a best practice here. Specific classes of potential solutions include reparameterization via theoretical means onto a linear scale (e.g., Wagenmakers, Krypotos, Criss, & Iverson, 2012), alternative statistical approaches (like conjoint measurement; Krantz & Tversky, 1971), experimental procedures that seek disordinal interactions, alternative experimental approaches (such as state-trace analysis; Loftus, Oberg, & Dillon, 2004; Newell & Dunn, 2008), and demonstrations of approximate linearity in the original measurement scale. In general, I would like to see more thoughtful application of inferential tests, catered to the data and specific question under investigation and less mindless and ritualistic (Gigerenzer, 2004; Loftus, 1996) application of traditional tools (something Tom Wickens used to call "the tyranny of the factorial ANOVA"). A good starting place for understanding current standards in reporting statistical tests is APA's *Journal Article Reporting Standards* (https://www.apastyle.org/jars).

Registered Reports

Registered reports are a type of article in which much of the review process and editorial decisions occur prior to data collection. The goal of this editorial policy is to select papers that are valuable by virtue of their rigorous methodology rather than exciting results. You can learn about the details of the procedure and the benefits at https://cos.io/rr/. I want to emphasize here that this is an appropriate avenue for publishing work for which any outcome, including a null result, is informative. Not every research endeavor fits this description.

Article Types

Other than Registered Reports, there will no longer be different article types in JEP:LMC. Sections of the journal indicating brief reports, observations, or commentary were appropriate in an era when issues of journals were read—or, at least, browsed. Now articles are consumed almost entirely individually, online, with no journal context. Obviously JEP:LMC is still open to any article that advances understanding of cognition, including experimental, meta-analytic, historical, and theoretical approaches. We also welcome commentary on previous articles and sometimes even on articles from other journals. Every submission will be evaluated not for its family resemblance to the historically prototypical JEP:LMC paper but for its unique contribution to a scientific understanding of cognition.

Article Preparation and Figures

I like APA style as much as the next editor. But please feel free to include tables and figures within the body of the text. It makes life easier for online reading. Just make sure to make them large enough and of sufficient precision for enlargement and close scrutiny.

Reviewing and Responding to Reviewers

I am extremely thankful to the many reviewers who contribute their time and expertise to evaluating articles for JEP:LMC. Let me gently remind you that your role is to evaluate the soundness and impact of the articles we send you and to provide suggestions for enhancing those qualities—but not to mold the paper in your image. Likewise, I would like to remind authors to treat reviewers' comments graciously, for they are, ultimately, free advice. Recognizing the value of that advice often requires an acknowledgment of a different perspective. In the current era, our treatment of one another reminds me of George Carlin's assessment of driving habits: Anyone driving slower than you is an idiot, and anyone going faster than you is a maniac. We are all sometimes insufficiently appreciative of reviewers and their labor when we are authors and are insufficiently respectful of authors' autonomy when we are reviewers.

I suspect that everyone will find something to grumble at about these policies. I do not imagine that the reformers will be entirely pleased with the fact that many choices remain voluntary in this scheme. Similarly, people who were happy with the way things were are likely to see any editorial changes as a personal burden. But if you were already doing high-quality, strongly motivated,

replicable, and reproducible research, then little about your life needs to change.

None of this attention to methodological and analytic detail means that theoretical impact will be devalued. We will continue to publish only creative, important research that advances the way we think about cognition. But it is my hope that, with only minor adjustments, we should expect well more than 50% of the results in this journal to convincingly replicate (cf. Open Science Collaboration, 2015).

JEP:LMC can remain a standard bearer for high-quality research in cognitive psychology in these heady times. The journal has successfully weathered our collective exuberance over splashy findings and the consequent turn to shorter and shorter articles. With your help, JEP:LMC can become the home for science that both can be replicated and is worthy of it.

References

Bartlett, F. C. (1932). *Remembering: An experimental and social study*. United Kingdom: Cambridge University.

Broadbent, D. E. (1980). The minimization of models. In A. J. Chapman & D. M. Jones (Eds.), *Models of man* (pp. 113–128). Leicester, England: British Psychological Society.

Gelman, A., & Loken, E. (2013). *The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time*. Unpublished paper. Retrieved from http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf

Gigerenzer, G. (2004). Mindless statistics. *Journal of Socio-Economics, 33,* 587–606. http://dx.doi.org/10.1016/j.socec.2004.09.033

Krantz, D. H., & Tversky, A. (1971). Conjoint-measurement analysis of composition rules in psychology. *Psychological Review, 78,* 151–169. http://dx.doi.org/10.1037/h0030637

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition, 6,* 312–319. http://dx.doi.org/10.3758/BF03197461

Loftus, G. R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science, 5,* 161–171. http://dx.doi.org/10.1111/1467-8721.ep11512376

Loftus, G. R., Oberg, M. A., & Dillon, A. M. (2004). Linear theory, dimensional theory, and the face-inversion effect. *Psychological Review, 111,* 835–863. http://dx.doi.org/10.1037/0033-295X.111.4.835

Neisser, U. (1976). *Cognition and reality*. New York, NY: W. H. Freeman and Company.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). New York, NY: Academic Press. http://dx.doi.org/10.1016/B978-0-12-170150-5.50012-3

Newell, B. R., & Dunn, J. C. (2008). Dimensions in data: Testing psychological models using state-trace analysis. *Trends in Cognitive Sciences, 12,* 285–290. http://dx.doi.org/10.1016/j.tics.2008.04.009

Open Science Collaboration. (2015). Psychology. Estimating the reproducibility of psychological science. *Science, 349,* aac4716. http://dx.doi.org/10.1126/science.aac4716

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science, 7,* 528–530. http://dx.doi.org/10.1177/1745691612465253

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. http://dx.doi.org/10.1177/0956797611417632

Wagenmakers, E. J., Krypotos, A. M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition, 40,* 145–160. http://dx.doi.org/10.3758/s13421-011-0158-0