

COMMENTARY

Emphasis on Direct Replication as Misplaced Physics Envy:
Commentary on [Clarke et al. \(2026\)](#)Jolynn Pek¹, Leandre R. Fabrigar², Duane T. Wegener¹, and Richard E. Petty¹¹Department of Psychology, The Ohio State University²Department of Psychology, Queen's University

[Clarke et al. \(2026\)](#) estimated that direct replications constitute only 0.2% of published articles in the 100 highest impact psychology journals, highlighting their claimed centrality in improving psychological science. In this commentary, we argue that an exclusive focus on direct replication overlooks the value of a broader set of replication strategies for evaluating the reliability of empirical findings. In particular, we emphasize that exact conceptual replications, which engage traditional validity considerations, namely, construct, internal, and statistical conclusion, can more precisely reveal alternative explanations for nonreplication outcomes than direct replications alone. We further contend that [Clarke et al.](#)'s characterization of replication as a precursor to validity might be a misrepresentation. Instead, we explain why validity and replication are fundamentally intertwined: Replication outcomes are only informative to the extent that the underlying operationalizations, measurements, and designs possess adequate validity. Finally, we suggest that [Clarke et al.](#) might have overstated and oversimplified the role of direct replication in scientific self-correction. Strengthening validity—through improved measurement, design, and theoretical specification—often provides a more powerful and informative route to scientific self-correction than repeated direct replications alone.

Keywords: direct replication, exact conceptual replication, conceptual replication, psychometric invariance, validity

[Clarke et al. \(2026\)](#) emphasized the importance of direct replications, reporting that direct replications constitute 0.2% of published articles in the 100 highest impact psychology journals. We examine three of [Clarke et al.](#)'s claims, especially highlighting implications of psychometric invariance to motivate a broader consideration of replication.

**Claim: Direct Replications Are the Central Method for
Assessing Reliability**

[Clarke et al. \(2026\)](#) focused on direct replications (studies that duplicate the experimental and sampling procedures of

an original investigation). This focus implicitly assumes that findings are tightly bound to the specific operations used—an assumption more characteristic of the physical than social sciences ([Petty, 2018](#)). In physics, there is little or no distance between constructs and their operations. Psychologists might envy such simplicity, but doing so mischaracterizes the nature of most psychological findings.

Psychological findings generally concern underlying constructs, not the specific operations used to manipulate or measure them. Psychometrics draws a clear distinction between constructs and operations, optimizing operations to best capture the intended constructs for a population and context. For example, Grade 3 and 6 math items both reflect math ability but for different developmental levels; experiments are tailored to specific populations and contexts with the understanding that their operationalizations might not broadly generalize. When the same operation (method) instantiates a different construct in a direct replication, the original and replication results become incomparable—failed psychometric invariance ([Fabrigar & Wegener, 2016](#)). [Flake et al. \(2022\)](#) documented various construct-validity concerns

Joshua Clegg served as action editor.

Jolynn Pek  <https://orcid.org/0000-0002-9694-4967>

Duane T. Wegener  <https://orcid.org/0000-0001-7639-4251>

Richard E. Petty  <https://orcid.org/0000-0002-2870-8575>

Correspondence concerning this article should be addressed to Jolynn Pek, Department of Psychology, The Ohio State University, 1835 Neil Avenue, 225 Psychology Building, Columbus, OH 43210, United States. Email: pek.5@osu.edu

including measurement nonequivalence between original studies and direct replications, noting that a replication cannot be considered valid without measurement equivalence (a part of psychometric invariance). An intended direct replication by Ebersole et al. (2016) did not reproduce the original findings, but Luttrell et al. (2017) successfully replicated the original effect when they approximated an exact conceptual replication (Petty, 2018) as did another attempt by Ebersole et al. (2017). Exact conceptual replications focus on the conceptual variables rather than the original operations, thereby maintaining psychometric invariance with the original study while matching the original study's population and contextual characteristics as closely as possible.

In psychology, it would seem useful to distinguish between (a) the concept of replication as assessing whether a construct-level finding recurs, (b) the methods used to assess replication (e.g., conceptual, direct, exact conceptual), and (c) the empirical outcome of whether the finding recurs. Broadening the methods and contexts for replication implies that replication is underdocumented when focusing only on direct replication and only in articles wherein the primary purpose is replication. Many articles include internal direct and/or (exact) conceptual replications, and many articles focused on moderation or extension include replication cells.

Direct replications are most relevant when the operations themselves matter, which is more likely in applied research. When the goal is to evaluate construct-level effects, however, exact conceptual replications using psychometrically invariant operations that might differ from the original are more appropriate. Further, researchers might purposefully use different operations to test whether the same construct-level effects emerge, thereby strengthening construct validity by demonstrating that the intended construct, not the specific operations, drives the effect. One need not focus on only one type of replication as in Clarke et al. (2026), but if one had to champion a particular type of replication, exact conceptual replication (emphasizing the broad set of traditional validities) might represent a better choice for psychology than direct replication (see also Crandall & Sherman, 2016).

Claim: Replication Is a Precursor to Other Elements of Research Quality

Direct replications aimed at making claims about the operations might not require construct validity, but the other validities remain important for interpreting replication outcomes. Exact conceptual replication is essentially an assessment of external validity that is similarly dependent on construct, internal, and statistical conclusion validity. That is, a failed direct replication might reveal an unreliable effect of the desired psychological construct, but nonreplication outcomes can also stem from other validity-related causes.

We argue that construct validity is a key prerequisite for meaningful replicability assessments in basic psychological research. Because exact conceptual replications are characterized by psychometric invariance, they require careful evaluation of whether manipulations and measures adequately capture the intended constructs. Establishing psychometric invariance can more precisely reveal alternative explanations for nonreplication outcomes. Internal validity (ensuring the design supports causal inferences comparable to the original study) and statistical conclusion validity (avoiding Type I and Type II errors) are likewise essential. They ensure that claims about external validity (that construct relations generalize across samples, times, cultures, settings, etc.) are not confounded by threats to construct, internal, or statistical conclusion validities (see Cook & Campbell, 1979; Fabrigar et al., 2020).

Claim: Direct Replications Are Central to Facilitate Scientific Self-Correction

Direct replications, as defined by Clarke et al. (2026), are not necessarily the primary mechanism through which science self-corrects. Historically, major scientific shifts have occurred without failed replications. Kuhn's account of paradigm change emphasized that accumulated anomalies—not failed direct replications—erode consensus and motivate new frameworks. Even in physics, self-correction rarely involves failed direct replication. The Copernican model, for example, replaced the Ptolemaic system through increasingly precise astronomical measurements and the development of Kepler's laws, not through direct replications of Ptolemy's procedures. In psychology, phrenology faded when empirical evidence showed that its operationalizations failed to measure the constructs they purported to capture (e.g., animal lesion studies demonstrating that brain functions were not localized as phrenology proposed). These cases illustrate that scientific self-correction often arises from improved theory, measurement, and evidence—from strengthening validity—rather than from direct replication failures.

Scientific self-correction operates through many channels that enhance validity—commentary, debate, new methods, new theories, and new data. Replication studies might sometimes accelerate this process, but the field must still determine which replications are most informative. Not all corrections carry the same weight; some refine effect sizes, whereas others overturn core conceptual claims. Our aim has been to highlight how different replication strategies—beyond direct replication—vary in their focus on operations, constructs, and psychometric invariance, and thus yield different kinds of information about methods and theory. We hope this commentary encourages continued methodological development in evaluating and designing informative replication studies.

References

- Clarke, B., Lee, P. Y. K., Schiavone, S. R., Rhemtulla, M., & Vazire, S. (2026). The prevalence of direct replication articles in top-ranking psychology journals. *American Psychologist*, 81(2), 238–253. <https://doi.org/10.1037/amp0001385>
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Rand McNally.
- Crandall, C. S., & Sherman, J. W. (2016). On the scientific superiority of conceptual replications for scientific progress. *Journal of Experimental Social Psychology*, 66, 93–99. <https://doi.org/10.1016/j.jesp.2015.10.002>
- Ebersole, C. R., Alaei, R., Atherton, O. E., Bernstein, M. J., Brown, M., Chartier, C. R., Chung, L. Y., Hermann, A. D., Joy-Gaba, J. A., Line, M. J., Rule, N. O., Sacco, D. F., Vaughn, L. A., & Nosek, B. A. (2017). Observe, hypothesize, test, repeat: Luttrell, Petty and Xu (2017) demonstrate good science. *Journal of Experimental Social Psychology*, 69, 184–186. <https://doi.org/10.1016/j.jesp.2016.12.005>
- Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., Baranski, E., Bernstein, M. J., Bonfiglio, D. B. V., Boucher, L., Brown, E. R., Budiman, N. I., Cairo, A. H., Capaldi, C. A., Chartier, C. R., Chung, J. M., Cicero, D. C., Coleman, J. A., Conway, J. G., ... Nosek, B. A. (2016). Many labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, 67, 68–82. <https://doi.org/10.1016/j.jesp.2015.10.012>
- Fabrigar, L. R., & Wegener, D. T. (2016). Conceptualizing and evaluating the replication of research results. *Journal of Experimental Social Psychology*, 66, 68–80. <https://doi.org/10.1016/j.jesp.2015.07.009>
- Fabrigar, L. R., Wegener, D. T., & Petty, R. E. (2020). A validity-based framework for understanding replication in psychology. *Personality and Social Psychology Review*, 24(4), 316–344. <https://doi.org/10.1177/1088868320931366>
- Flake, J. K., Davidson, I. J., Wong, O., & Pek, J. (2022). Construct validity and the validity of replication studies: A systematic review. *American Psychologist*, 77(4), 576–588. <https://doi.org/10.1037/amp0001006>
- Luttrell, A., Petty, R. E., & Xu, M. (2017). Replicating and fixing failed replications: The case of need for cognition and argument quality. *Journal of Experimental Social Psychology*, 69, 178–183. <https://doi.org/10.1016/j.jesp.2016.09.006>
- Petty, R. E. (2018). The importance of exact conceptual replications. *Behavioral and Brain Sciences*, 41, Article e146. <https://doi.org/10.1017/S0140525X18000821>

Received May 11, 2026

Revision received August 27, 2026

Accepted September 4, 2026 ■