

Rethinking the Evaluation of Psychological Theories: A Commentary on Calderon et al.'s Registered Replication Report

Ihnwhi Heo and Minjeong Jeon

University of California, Merced, Merced, California

Abstract

Construal-level theory (CLT) is a well-established theory in social psychology that posits a relationship between psychological distance and mental abstraction such that greater distance is associated with more abstract representations. In a registered replication report, Calderon et al. presented findings that call into question the strength and generality of this relationship, including effects that were attenuated, null, or even opposite to those predicted by CLT across multiple domains of psychological distance. In this commentary, we engage with these findings at two complementary levels. We first reflect on design and measurement considerations that shape the evidentiary value of the current replication. We then situate the findings in discussions of theory evaluation amid ongoing discourse on the theory crisis in psychology. Rather than framing the results in terms of theoretical success or failure, we invite researchers to reconsider how psychological theories are evaluated, revised, and sustained as evidence accumulates. We outline how a findable, accessible, interoperable, and reusable (FAIR)-informed approach, applied to CLT, may support more transparent and cumulative theory development.

Keywords

construal-level theory, cumulative science, FAIR theory, replication, theory crisis, theory evaluation

Received 12/29/25; revision accepted 3/24/26

Human cognition is not confined to the immediate here and now. People routinely anticipate future events, imagine distant places, consider the perspectives of socially removed others, and reflect on events that are uncertain or hypothetical. Construal-level theory (CLT; Liberman & Trope, 1998; Trope & Liberman, 2003) proposes an account of how individuals mentally represent experiences that extend beyond direct, immediate engagement. At its core, CLT centers on the concept of psychological distance, which may vary along several dimensions: temporal, spatial, social, and likelihood (or hypotheticality). A central claim of the theory is that as events become more distant along any of these dimensions, they are represented at higher levels of abstraction, emphasizing broader and more general features rather than concrete, contextual details. Conversely, psychologically proximal

events are expected to be represented in more concrete and detail-oriented terms. In this way, CLT proposes a relationship between psychological distance and mental abstraction such that greater distance corresponds to more abstract mental representations.

Over the past two decades, this claim has become highly influential, generating extensive empirical research across social and behavioral domains beyond psychology. A growing number of replication efforts have reported mixed evidence, including attenuated effect sizes relative to earlier findings and in some cases,

Corresponding Author:

Ihnwhi Heo, Department of Psychological Sciences, University of California, Merced, Merced, California
Email: ihnwhi.heo@gmail.com



null results (e.g., Calderon et al., 2020; Gong & Medin, 2012; Luke et al., 2021; McCarthy et al., 2018; Sánchez et al., 2021; Žeželj & Jokić, 2014). Following mixed and attenuated findings, Calderon et al. (2025) conducted a large-scale, multilab registered replication report (RRR) to more rigorously test the proposed relationship between psychological distance and mental abstraction. They evaluated four core domains of psychological distance—temporal, spatial, social, and likelihood—and examined whether each increased levels of abstraction. Temporal and spatial distance were assessed through close replications of classic paradigms introduced by Liberman and Trope (1998) and Fujita et al. (2006), and social distance and likelihood were tested through paradigmatic replications intended to capture the same theoretical constructs. Across these four domains, the results revealed weak, null, or even directionally opposite effects, raising questions about the robustness and generality of the proposed relationship between psychological distance and abstraction.

Given its scale, transparency, and methodological rigor, this RRR of Calderon et al. (2025) represents an important contribution to the CLT literature. At the same time, findings of this nature invite careful reflection on how evidence from large, well-powered replications should be interpreted in relation to established psychological theories. Do such results warrant strong conclusions about theoretical invalidity, or do they instead call for closer examination of operationalizations, design features, and the scope conditions under which a theory is expected to hold? Regardless of how these results are interpreted, how should they be incorporated into theory development? In this commentary, we engage with these questions at two complementary levels. We first consider how design and measurement choices may shape the evidentiary value of the present replication. We then situate the findings in broader discussions of theory evaluation in psychology, including recent discourse on the theory crisis.

Design and Measurement Considerations in Calderon et al.

Calderon et al. (2025) reported findings indicating that the relationship between psychological distance and mental abstraction as it has been commonly understood in the literature is not supported. Identifying such nonsupportive evidence through replication efforts is both informative and consequential. The theoretical implications of such evidence are not self-evident but depend on how the constructs were operationalized and tested and the scope of inference afforded by the chosen designs. We therefore begin by considering how the findings reported by Calderon et al. should be

interpreted for the evaluation of CLT. In particular, we focus on the design and measurement choices that structure what such large-scale, multilab replications can and cannot reveal.

From a design perspective, Calderon et al. (2025) employed direct replications for temporal and spatial distance and paradigmatic replications for social distance and likelihood. These design choices closely reproduce or are explicitly grounded in the original paradigms introduced by Liberman and Trope (1998) and Fujita et al. (2006). This approach strengthens the interpretability of the findings with respect to the original studies and enables a clear and principled assessment of replicability. At the same time, reliance on the original design and measurement choices may have constrained the range of theoretical inferences regarding broader claims about the operation and boundaries of CLT beyond these specific paradigmatic instantiations.

Design features related to experimental manipulation further inform the interpretation of Calderon et al.'s (2025) findings. The between-subjects priming approach employed throughout the replication, in which each participant was exposed to only one level of psychological distance (e.g., “tomorrow” [close condition] vs. “next year” [distant condition] for temporal distance), has been shown to yield relatively modest and unstable effects in large-scale replication efforts (Camerer et al., 2016). Complementary evidence suggests that likelihood manipulations tend to be more effective when both high- and low-likelihood conditions are jointly presented because this allows participants to process the intended contrast more explicitly (Grinfeld et al., 2024).

These design considerations point to the potential role of contextual and methodological moderators in shaping observed CLT effects. Both Soderberg et al. (2015) and Maier et al. (2022) documented substantial heterogeneity across studies, indicating that CLT effects are neither uniform nor invariant across operationalizations. As Maier et al. emphasized, progress in this literature may benefit from analytic and design strategies that explicitly model such heterogeneity rather than treating it as residual noise. Viewed in this light, Calderon et al. (2025) provided evidence about the performance of specific paradigms under large-scale conditions while leaving open how CLT operates across a range of theoretically relevant contexts.

Turning to measurement considerations, a substantial body of prior work suggests that the magnitude and consistency of CLT effects are sensitive to how psychological distance and mental abstraction are operationalized. In their meta-analysis, Soderberg et al. (2015) showed that effect sizes varied systematically as a function of temporal-distance implementations such that larger effects were observed when temporal contrasts

were positioned further into the future. They also discussed participant engagement as a possible factor contributing to variation in effects across distance manipulations. From this perspective, it is noteworthy that Calderon et al. (2025) operationalized temporal proximity using “tomorrow” as the near condition and employed imagined distance manipulations involving limited elaboration (e.g., presenting near vs. distant time points or locations without extended imaginative tasks), choices that may be associated with attenuated temporal-distance effects. Under such conditions, attenuated findings may be more informatively interpreted as reflecting the consequences of particular operational choices rather than as decisive evidence against the underlying theoretical mechanism.

Accordingly, evaluations of CLT that focus exclusively on direct or paradigmatic replications of its earliest paradigms may, by necessity, offer a narrower window onto the theory’s empirical status. Incorporating design and measurement approaches that have yielded more robust effects in prior work could allow for a more comprehensive assessment of the scope and limits of CLT. Importantly, if such theoretically favorable conditions were to produce null, attenuated, or even directionally opposite effects, this would provide even stronger grounds for revisiting and refining the theory’s core claims.

Theory Evaluation Amid Theory Crisis

Beyond their empirical contributions, Calderon et al. (2025) raised questions about theory evaluation: that is, how evidence that fails to support a theory should be interpreted in relation to the theory itself and how divergent empirical findings should be integrated into a cumulative theoretical framework. These questions resonate with recent discussions of the theory crisis in psychology (Eronen & Bringmann, 2021; Muthukrishna & Henrich, 2019; Oberauer & Lewandowsky, 2019; Sanbonmatsu et al., 2025; van Lissa et al., 2026). Concerns about the theory crisis center on the claim that psychological theories tend to be just proposed and then abandoned without sustained development (Meehl, 1978). Such concerns have emerged against the backdrop of advances in psychological-research practices following the identification of the replication crisis (Open Science Collaboration, 2015) and the subsequent dissemination of open-science practices (Pfadt et al., 2025). From this perspective, replication efforts prompt renewed attention to the evidentiary standards by which psychological theories are sustained, revised, or reconsidered.

We believe that ongoing methodological developments in psychological science will intensify scrutiny of existing theories and reshape the standards by which they are evaluated. The field’s capacity to detect smaller

effects and previously unappreciated sources of variability has increased in recent years, driven by advances in big data for psychological research (Adjerid & Kelley, 2018; Harlow & Oswald, 2016), the growth of multilab replication collaborations (Baumeister et al., 2023; Stroebe, 2019), paradigm shifts associated with machine-learning approaches (Dwyer et al., 2018; Yarkoni & Westfall, 2017), and the increasing use of multimodel inference to represent multiple competing theoretical accounts (Heo et al., 2025; Hinne et al., 2020). As a consequence, psychological theories are likely to encounter more frequent instances of apparent nonreplication or reduced effect sizes even when underlying mechanisms operate only under restricted conditions. In this evolving methodological landscape, the evaluation of theories becomes increasingly nuanced. Treating apparent nonreplications or reduced effect sizes as definitive refutations risks conflating limitations of empirical tests with limitations of the theory itself, whereas dismissing these findings risks insulating theories from rigorous evaluation. How this balance should be struck remains an open question in the field.

In this context, the findings of Calderon et al. (2025) provide a concrete case through which to examine how established theories, such as CLT, ought to be evaluated in light of challenging evidence. CLT has long occupied a central position in social psychology. When such a theory is challenged by evidence from a highly powered, multilab RRR as in Calderon et al., the question is not simply whether the theory passes or fails but how the field should interpret and respond to such challenges. For CLT in particular, the more constructive question may be how the theory might be refined, clarified, or bounded to enhance its practical applicability. More generally, how far should theoretical clarification and qualification extend when empirical support remains mixed? How should theory development balance revision, boundary specification, and integration as evidence accumulates? And how should psychologists navigate the tension between sustaining established theoretical frameworks and engaging with disconfirming or ambiguous findings in ways that promote cumulative understanding?

Reconsidering CLT in light of challenging evidence

The findings of Calderon et al. (2025) do give us pause, although not in ways that imply uniform disconfirmation across all components of the theory. The most consequential challenge may concern CLT’s claim that psychological distance constitutes a unified construct whose effects on abstraction are equivalent across temporal, spatial, social, and likelihood dimensions. The pattern

of results reported by Calderon et al., spanning weak, null, and directionally opposite effects across domains, raises the possibility that these four dimensions do not function as interchangeable instantiations of a common mechanism. If so, the theoretical architecture of CLT, which treats psychological distance as a unitary construct, requires closer examination. Whether this cross-domain divergence reflects differences in the nature of the constructs themselves, in how they engage motivational systems, or in the boundary conditions under which abstraction is triggered remains to be clarified. But the inconsistency across domains is not easily attributable to operationalization alone and constitutes a signal for theoretical revision.

Another aspect that warrants reconsideration concerns the relationship between effect magnitudes and the applied scope of CLT. The theory has been invoked to explain and predict behavior across multiple domains beyond psychology. The practical relevance of these extensions depends on both whether the abstraction-distance link exists and whether it is sufficiently robust and consistent to support the downstream predictions built on it. If effects are small, highly context-dependent, and sensitive to operationalization, as Calderon et al.'s (2025) findings might suggest, the implications for applied domains require careful qualification. The findings of Calderon et al. thus invite reflection on the existence of the core mechanism, the scope and magnitude of effects that can be reliably expected, and the conditions under which the theory's applied extensions remain justified.

Toward cumulative theory development: A FAIR-informed approach

One way to approach these questions is to reconsider not only whether CLT is supported but also how the theory itself is specified, archived, and revised as evidence accumulates. If the relationship between psychological distance and abstraction yields attenuated or context-dependent effects, this need not immediately imply theoretical invalidity. It may instead reveal ambiguities in what the theory commits to under different operational conditions—how core constructs are defined, how boundary conditions are articulated, and how auxiliary assumptions are incorporated. The challenge, then, is not merely empirical but also structural: Psychological theories are often articulated in prose, with limited formal specification of scope, dependencies, and revision rules. In such cases, mixed findings accumulate without a clear framework for determining when revision, qualification, or reconstruction is warranted.

Recent discussions of findable, accessible, interoperable, and reusable (FAIR) theory offer one potential way to address this structural problem (Lamprecht et al., 2020; van Lissa et al., 2026; Wilkinson et al., 2016). The FAIR framework extends the principles of making scientific objects findable, accessible, interoperable, and reusable to the domain of theory itself. Rather than treating a theory as a static textual claim, the FAIR framework conceptualizes theory as a versioned, digital object that can be explicitly specified, archived, and iteratively refined. For a concrete procedural workflow, see van Lissa et al. (2026), who outlined steps through the R *theorytools* package for specifying, versioning, and archiving theories as digital objects. Building on this framework, we illustrate how a FAIR-informed approach can be applied to CLT to address the questions raised above in a procedurally explicit manner.

Applied to CLT, a FAIR-informed approach would not entail formalizing every aspect of the theory at once or adopting a particular mathematical representation. Instead, it would involve distinguishing between the theory's core propositions and the specific paradigmatic instantiations through which those propositions have been tested and making these components explicitly linkable to empirical evidence. Under such a structure, null or attenuated replication findings, such as those reported by Calderon et al. (2025), could be mapped onto particular operationalizations (e.g., specific magnitudes of distance contrasts, measurement strategies, or contextual moderators) rather than being interpreted as wholesale confirmation or disconfirmation of the theory as a whole.

When empirical support remains mixed, theoretical clarification need not take the form of ad hoc adjustment. In a FAIR-informed framework, revisions and boundary specifications become documented and traceable across versions rather than implicit adjustments made in response to isolated findings. Version control makes it possible to differentiate between modest boundary refinement (e.g., specifying that abstraction shifts occur only under sufficiently large psychological distance contrasts), auxiliary modification (e.g., incorporating motivational or contextual moderators), and more substantive reconstruction of the theory's explanatory structure. This traceability reduces the risk of strategic ambiguity (Eisenberg, 1984; Frankenhuis et al., 2023) while preserving space for principled theoretical development and more explicit theory formalization.

Balancing revision, boundary specification, and integration as evidence accumulates likewise becomes more systematic under FAIR principles. Because theories are treated as identifiable and versioned objects, empirical findings can be linked to particular theoretical components. Domain-specific claims—for instance, whether

CLT operates equivalently across temporal, spatial, social, and likelihood dimensions—can be specified modularly and evaluated comparatively. Competing refinements, including stronger or weaker formulations of the abstraction-distance link, may coexist transparently rather than displacing one another implicitly. In this way, theoretical evolution becomes structured rather than diffuse, and integration across findings becomes cumulative rather than piecemeal.

Importantly, adopting FAIR principles does not insulate theories from refutation. On the contrary, it may heighten theoretical accountability by requiring clearer articulation of what would count as support, what constitutes revision, and what would warrant abandonment. For CLT, this could mean specifying more precisely the magnitude and form of psychological-distance effects, the domains across which generalization is expected, and the conditions under which the abstraction-distance relationship should attenuate or disappear. As evidence accumulates, such specifications could be updated in a transparent and traceable manner. In this sense, the findings of Calderon et al. (2025) may be viewed less as a verdict on CLT and more as an opportunity to reconsider how established theories are evaluated. Embedding theory development in a FAIR infrastructure does not resolve substantive disagreements, but it provides a framework for navigating the tension between sustaining influential frameworks and engaging with disconfirming or ambiguous findings in a way that promotes cumulative understanding.

Final Reflections

In the present commentary, we engaged with the findings of Calderon et al. (2025) at two complementary levels. Methodologically, we have emphasized how design and measurement choices shape the evidentiary value of large-scale replications and delimit the conclusions that can be drawn about a well-established theory such as CLT. Conceptually, we have used CLT as a concrete case to reflect on broader questions of theory evaluation, revision, and development. We have illustrated how a FAIR-informed perspective can render theoretical refinement more explicit, traceable, and cumulative.

Do the findings of Calderon et al. (2025) warrant reconsideration of CLT? In our view, they do—but not in a manner that implies wholesale theoretical abandonment. Rather, the results invite closer scrutiny of the scope, strength, and boundary conditions of the relationship between psychological distance and abstraction. Because highly powered replications yielded attenuated, null, or even effects opposite to those predicted by CLT under paradigms aligned with foundational studies, this raises legitimate questions about robustness and

generalizability. At the same time, such findings more plausibly motivate clearer specification of expected effect magnitudes, operationalizations, and contextual moderators than categorical rejection of the theory's core conceptual insight.

From a FAIR-informed perspective, this process can be understood not as a binary verdict on the validity of CLT but as an opportunity to make the theory's assumptions, operationalizations, and evidentiary links more explicit and systematically revisable. In this sense, the findings of Calderon et al. (2025) exemplify how replication evidence can contribute to cumulative theory development by clarifying which components of a theory require refinement, boundary specification, or further empirical examination. As psychological science continues to advance methodologically, established theories will likely face increasingly visible challenges. Embedding theory development in transparent and cumulative infrastructures may therefore be less a reaction to the theory crisis than a necessary step in sustaining principled theoretical progress.

Transparency

Action Editor: David A. Sbarra

Editor: David A. Sbarra

Author Contributions

Ihnwhi Heo: Conceptualization; Project administration; Writing – original draft; Writing – review & editing.

Minjeong Jeon: Conceptualization; Writing – original draft; Writing – review & editing.

Declaration of Conflicting Interests

The authors declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

ORCID iD

Ihnwhi Heo  <https://orcid.org/0000-0002-6123-3639>

References

- Adjerid, I., & Kelley, K. (2018). Big data in psychology: A framework for research advancement. *American Psychologist*, 73(7), 899–917. <https://doi.org/10.1037/amp0000190>
- Baumeister, R. F., Tice, D. M., & Bushman, B. J. (2023). A review of multisite replication projects in social psychology: Is it viable to sustain any confidence in social psychology's knowledge base? *Perspectives on Psychological Science*, 18(4), 912–935. <https://doi.org/10.1177/17456916221121815>
- Calderon, S., Mac Giolla, E., Ask, K., Adler, S. J., Agerström, J., Akpinar, B., Albayrak, N., Alparone, F. R., Amin, S., Aquino, A., Bachet, M., Baisile, B., Bausenhardt, K. M., Beylat, M., Bialobrzaska, O., Bloomfield, E. C., Boecker, L., Bonora, M., Brady, S. T., . . . Luke, T. J. (2025). *Effects of psychological distance on mental abstraction: A registered report of four tests of construal level theory*. PsyArXiv. https://doi.org/10.31234/osf.io/wqbhd_v2

- Calderon, S., Mac Giolla, E., Ask, K., & Granhag, P. A. (2020). Subjective likelihood and the construal level of future events: A replication study of Wakslak, Trope, Liberman, and Alony (2006). *Journal of Personality and Social Psychology*, 119(5), e27–e37. <https://doi.org/10.1037/pspa0000214>
- Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmeld, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280), 1433–1436. <https://doi.org/10.1126/science.aaf0918>
- Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine learning approaches for clinical psychology and psychiatry. *Annual Review of Clinical Psychology*, 14, 91–118. <https://doi.org/10.1146/annurev-clinpsy-032816-045037>
- Eisenberg, E. M. (1984). Ambiguity as strategy in organizational communication. *Communication Monographs*, 51(3), 227–242. <https://doi.org/10.1080/03637758409390197>
- Eronen, M. I., & Bringmann, L. F. (2021). The theory crisis in psychology: How to move forward. *Perspectives on Psychological Science*, 16(4), 779–788. <https://doi.org/10.1177/1745691620970586>
- Frankenhuis, W. E., Panchanathan, K., & Smaldino, P. E. (2023). Strategic ambiguity in the social sciences. *Social Psychological Bulletin*, 18, 1–25. <https://doi.org/10.32872/spb.9923>
- Fujita, K., Trope, Y., Liberman, N., & Levin-Sagi, M. (2006). Construal levels and self-control. *Journal of Personality and Social Psychology*, 90(3), 351–367. <https://doi.org/10.1037/0022-3514.90.3.351>
- Gong, H., & Medin, D. L. (2012). Construal levels and moral judgment: Some complications. *Judgment and Decision Making*, 7(5), 628–638. <https://doi.org/10.1017/S1930297500006343>
- Grinfeld, G., Wakslak, C., Trope, Y., & Liberman, N. (2024). Construing hypotheticals: How hypotheticality affects level of abstraction. *Journal of Experimental Social Psychology*, 110, Article 104543. <https://doi.org/10.1016/j.jesp.2023.104543>
- Harlow, L. L., & Oswald, F. L. (2016). Big data in psychology: Introduction to the special issue. *Psychological Methods*, 21(4), 447–457. <https://doi.org/10.1037/met0000120>
- Heo, I., Simons, J.-W., & Liu, H. (2025). A tutorial on Bayesian model averaging for exponential random graph models. *British Journal of Mathematical and Statistical Psychology*. Advance online publication. <https://doi.org/10.1111/bmsp.70007>
- Hinne, M., Gronau, Q. F., van den Bergh, D., & Wagenmakers, E.-J. (2020). A conceptual introduction to Bayesian model averaging. *Advances in Methods and Practices in Psychological Science*, 3(2), 200–215. <https://doi.org/10.1177/2515245919898657>
- Lamprecht, A.-L., Garcia, L., Kuzak, M., Martinez, C., Arcila, R., Martin Del Pico, E., Dominguez Del Angel, V., van de Sandt, S., Ison, J., Martinez, P. A., McQuilton, P., Valencia, A., Harrow, J., Psomopoulos, F., Gelpi, J. L., Hong, N. C., Goble, C., & Capella-Gutierrez, S. (2020). Towards fair principles for research software. *Data Science*, 3(1), 37–59. <https://doi.org/10.3233/DS-190026>
- Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, 75(1), 5–18. <https://doi.org/10.1037/0022-3514.75.1.5>
- Luke, T. J., Ask, K., Magnusson, E., Calderon, S., & Mac Giolla, E. (2021). *Revisiting the relationship between social distance and communication preferences: Replications and reinterpretation of Amit et al. (2013, Experiment 2)*. PsyArXiv. <https://doi.org/10.31234/osf.io/qjw7g>
- Maier, M., Bartoš, F., Oh, M., Wagenmakers, E.-J., Shanks, D., & Harris, A. J. (2022). *Adjusting for publication bias reveals that evidence for and size of construal level theory effects is substantially overestimated*. PsyArXiv. <https://doi.org/10.31234/osf.io/r8nyu>
- McCarthy, R. J., Hartnett, J. L., Heider, J. D., Scherer, C. R., Wood, S. E., Nichols, A. L., Edlund, J. E., & Walker, W. R. (2018). An investigation of abstract construal on impression formation: A multi-lab replication of McCarthy and Skowronski (2011). *International Review of Social Psychology*, 31(1), Article 15. <https://doi.org/10.5334/irsp.133>
- Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46(4), 806–834. <https://doi.org/10.1037/0022-006X.46.4.806>
- Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour*, 3(3), 221–229. <https://doi.org/10.1038/s41562-018-0522-1>
- Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, 26, 1596–1618. <https://doi.org/10.3758/s13423-019-01645-2>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), Article aac4716. <https://doi.org/10.1126/science.aac4716>
- Pfadt, J. M., Bartoš, F., Godmann, H. R., Waaijers, M., Groot, L., Heo, I., Mensink, L., Nak, J., de Ruiter, J. P., Sarafoglou, A., Siepe, B., Arena, G., Akrong, E., Aust, F., van den Bergh, D., Brenner, W., Doekemeijer, R., Donzallaz, M. C., van Doorn, J., . . . Wagenmakers, E.-J. (2025). *A methodological metamorphosis: The rapid rise of Bayesian inference and open science practices in psychology*. PsyArXiv. https://doi.org/10.31234/osf.io/ck3js_v1
- Sanbonmatsu, D. M., Neufeld, B., & Posavac, S. S. (2025). There is no theory crisis in psychological science. *Journal of Theoretical and Philosophical Psychology*. Advance online publication. <https://doi.org/10.1037/teo0000301>
- Sánchez, A. M., Coleman, C. W., & Ledgerwood, A. (2021). Does temporal distance influence abstraction? A large pre-registered experiment. *Social Cognition*, 39(3), 352–365. <https://doi.org/10.1521/soco.2021.39.3.352>
- Soderberg, C. K., Callahan, S. P., Kochersberger, A. O., Amit, E., & Ledgerwood, A. (2015). The effects of psychological distance on abstraction: Two meta-analyses. *Psychological Bulletin*, 141(3), 525–548. <https://doi.org/10.1037/bul0000005>

- Stroebe, W. (2019). What can we learn from many labs replications? *Basic and Applied Social Psychology*, 41(2), 91–103. <https://doi.org/10.1080/01973533.2019.1577736>
- Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, 110(3), 403–421. <https://doi.org/10.1037/0033-295X.110.3.403>
- van Lissa, C. J., Peikert, A., Ernst, M. S., van Dongen, N. N., Schönbrodt, F. D., & Brandmaier, A. M. (2026). To be FAIR: Theory specification needs an update. *Perspectives on Psychological Science*, 21(2), 173–191. <https://doi.org/10.1177/17456916251401850>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., . . . Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, Article 160018. <https://doi.org/10.1038/sdata.2016.18>
- Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, 12(6), 1100–1122. <https://doi.org/10.1177/1745691617693393>
- Žeželj, I. L., & Jokić, B. R. (2014). Replication of experiments evaluating impact of psychological distance on moral judgment. *Social Psychology*, 45(3), 223–231. <https://doi.org/10.1027/1864-9335/a000188>