

## A Response to a Comment on Hall et al. (2024)

**Kathleen Schmidt<sup>1</sup> , Gerald J. Haefel<sup>2</sup> , Neil Levy<sup>3,4</sup> , David Moreau<sup>5,6</sup> , Sean T. H. Lee<sup>7</sup> , Erin M. Buchanan<sup>8</sup> , Anthony J. Krafnick<sup>9</sup> , Martin Voracek<sup>10</sup> , Gerit Pfuhl<sup>11</sup>, Krystian Barzykowski<sup>12</sup>, Marta Kowal<sup>13</sup> , and Jordan Wagge<sup>14</sup> **

<sup>1</sup>Department of Psychology, Ashland University, Ashland, Ohio; <sup>2</sup>Department of Psychology, University of Notre Dame, Notre Dame, Indiana; <sup>3</sup>Department of Philosophy, Macquarie University, Sydney, Australia; <sup>4</sup>Uehiro Centre for Practical Ethics, University of Oxford, Oxford, England; <sup>5</sup>School of Psychology, University of Auckland, Auckland, New Zealand; <sup>6</sup>Centre for Brain Research, University of Auckland, Auckland, New Zealand; <sup>7</sup>School of Social and Health Sciences, James Cook University, Singapore; <sup>8</sup>Analytics, Harrisburg University of Science and Technology, Harrisburg, Pennsylvania; <sup>9</sup>Department of Psychology, Dominican University, River Forest, Illinois; <sup>10</sup>Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria; <sup>11</sup>Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway; <sup>12</sup>Institute of Psychology, Faculty of Philosophy, Jagiellonian University, Krakow, Poland; <sup>13</sup>IDN Being Human Lab - Institute of Psychology, University of Wrocław, Wrocław, Poland; and <sup>14</sup>School of Psychology, Avila University, Kansas City, Missouri

### Keywords

epistemic intuitions, Gettier cases, knowledge, replication

Received 5/31/24; Revision accepted 7/1/24

The conditions under which laypeople attribute knowledge to protagonists have long been debated by experimental philosophers (Colaço et al., 2014; Nagel et al., 2013; Weinberg et al., 2001). Consider the case of “Darrel,” who accurately recognizes the species of an animal in the woods even though it was the only one of its kind among many animals of an identical-looking species. Responses to this so-called “Gettier-type” case have been studied to examine whether laypeople consider luckily true beliefs as constituting actual knowledge. In their Experiment 1, Turri et al. (2015) compared knowledge attributions in this version of the “Darrel” case to a version in which his belief is a clear case of knowledge; they found no difference between these conditions and concluded that “a salient but failed threat to the truth of a judgment does not significantly affect whether it is viewed as knowledge” (p. 381). Hall et al. (2024) replicated and extended Experiment 1 of Turri et al., testing condition differences by using the Darrel case alongside two other counterfeit-object Gettier-type cases in a large multinational study. Hall et al. found that participants were less likely to attribute knowledge to the protagonists when beliefs were only luckily true (i.e., the Gettier

conditions) than when the truth of the beliefs was not under threat (i.e., the knowledge conditions). This significant condition difference reported by Hall et al. stands in contrast to the null result reported by Turri et al.

In their commentary on Hall et al., Buckwalter and Friedman (2024) claimed that the replication should have been interpreted as successful, argued that the researchers’ conclusions were incorrect, and implied that the replication effort was misguided. As a subset of contributors to the Hall et al. replication, we appreciate the opportunity to respond to their comment. Although we recognize the potential for disagreement in the interpretation of research results, Buckwalter and Friedman’s critique ignored several key features of the research, and many of their arguments and proposed interpretations were already addressed by Hall et al. (2024). In response to their comment, we (a) explain why Hall et al. did not replicate all of the original findings, (b) emphasize how

### Corresponding Author:

Jordan Wagge, School of Psychology, Avila University, Kansas City, Missouri

Email: jordan.wagge@avila.edu



Hall et al. were accurate and nuanced in the description and interpretation of their results, and (c) caution against focusing on the mechanisms underlying a psychological phenomenon before it is clearly established.

## Different Statistical and Interpretive Conclusions

Although Hall et al. (2024) found evidence for several of the original claims of Turri et al. (2015), they did not replicate one notable null result: Unlike Turri et al., they observed a difference in knowledge attribution between the Gettier condition and the knowledge control condition. Although small in magnitude, this difference was identified robustly across three tested vignettes, different measures of knowledge attribution, and multiple analyses. Hall et al. further examined possible moderators, none of which eliminated the condition difference that Turri et al. failed to find. Given two studies that differ on the statistical significance of their results, we contend that the study with the larger, more diverse sample and higher analytic power to detect differences is more likely to yield valid, replicable, and generalizable results. Because the difference between the aforementioned conditions is absolutely central to the philosophical debate over Gettier cases, and the experimental philosophical literature has accordingly focused on it, we are confident that Hall et al.'s failure to replicate the original null result is an important contribution.

Nonetheless, Buckwalter and Friedman (2024) concluded that the two discrepant findings were “basically the same.” This assertion is based on their interpretation of descriptive differences in the frequencies of knowledge attributions between conditions and enables them to generate a narrative (or, reframing) that would align with that of Turri et al. (2015). However, Hall et al.'s (2024) preregistered analysis plan did not include this method of assessing the results. The purpose of preregistration is to reduce researcher degrees of freedom and to discourage exploratory analyses disguised as confirmatory, especially those that support preferred narratives (e.g., Nosek et al., 2018; Wicherts et al., 2016). In this case, Hall et al. anticipated a null result aligned with the original finding, conducted a prespecified analysis with any data-driven changes transparently noted, detected a nominally significant result, and reported those findings accordingly. Based on these prespecified conditions and consistent with the framework of null hypothesis significance testing (NHST), we must state that Hall et al.'s findings were different from those of the original study.

Although one can, and Hall et al. (2024) did, calculate an effect size from the null result reported in Experiment 1 of Turri et al. (2015; i.e., odds ratio [OR] = 2.00, 95% confidence interval [CI] = [0.77, 5.21]), interpreting the

results of the statistical test beyond a failure to reject the null hypothesis would violate the logic of NHST (e.g., Frick, 1996). The null result merely indicates that Turri et al. had insufficient evidence to support a claim of a difference between conditions. Nonetheless, as Hall et al. acknowledged, the effect they observed (i.e., OR = 1.86, 95% CI = [1.78, 1.94]) was similar in size to the original. Accordingly, as both Hall et al. and Buckwalter and Friedman (2024) pointed out, the large sample in the replication study may be one explanation for the difference in statistical conclusions. Lacking this insight, Turri et al. concluded that knowledge attributions were insensitive to the differences in Gettier and knowledge conditions in their Experiment 1—a claim that was not supported by the results of Hall et al. Nonetheless, Buckwalter and Friedman suggested that the percentage differences in the original, likely underpowered, Turri et al. study should have been the sole criteria for replication rather than the statistical conclusions or interpretations therefrom. However, the CI of the original effect is quite wide and includes effect sizes that would imply a greater likelihood of knowledge attribution in the knowledge condition than the Gettier condition ( $OR > 1$ ), a null result ( $OR = 1$ ), and a greater likelihood of knowledge attribution in the Gettier condition than the knowledge condition ( $OR < 1$ ); therefore, their results were inconclusive. In contrast, the results from Hall et al. represent a clear, albeit small, Gettier-intuition effect. The results of the original study simply cannot be classified as “basically the same” as those presented by Hall et al.

## Accurate and Nuanced

Contrary to Buckwalter and Friedman's (2024) reading, Hall et al. (2024) did not “claim to . . . show that people deny knowledge to lucky agents” or “infer from this finding that there is a common psychological tendency to deny that true-by-luck beliefs are knowledge.” Rather, Hall et al. accurately reported a nominally significant condition difference, which they then interpreted accordingly. For example, they stated the following:

This result did not correspond to that found by Turri et al., who failed to detect a significant difference in knowledge attribution between these two conditions . . . we did find effect sizes in the same range as Turri et al. when directly comparing like conditions; however, the null result did not replicate.

Hall et al. (2024) further noted their sample's baseline skepticism (i.e., that a “notable number of participants [43.41%; see Table 7] denied knowledge to protagonists even in clear cases of justified true belief”) and for this

reason, were careful in their interpretation of the condition difference. Specifically, they acknowledged that “Gettier intuitions were by no means common,” “the small size of the effect suggests that Gettier intuitions were not prevalent in our research,” and that “given the small size of the observed effect, the theoretical significance of this result is debatable.” They also provided a nuanced discussion of how differences in methods, design, analytic approach, and geographical origin and composition of the samples could have affected results.

### Constructs and Mechanisms

We agree with Buckwalter and Friedman (2024) that researchers who claim to answer vague questions about poorly defined constructs may very well lead readers to inappropriate conclusions. However, Hall et al.’s (2024) rhetorical use of “Gettier-type case” and “Gettier intuition” did not preclude the specificity of their claims or the recognition of their constraints on generality. Their limitations were readily acknowledged. Contrary to what Buckwalter and Friedman implied, Hall et al. did not make strong statements about the theoretical implications of their results.

Buckwalter and Friedman (2024) argued for the development of theories and identification of mechanisms in research on luckily true beliefs. We agree that the study of mechanisms is important; in fact, Hall et al. (2024) explored the moderating role of luck attributions directly in their analyses.<sup>1</sup> However, the examination of mechanisms should, in our view, follow the confirmation of an effect’s replicability. The approach of Hall et al. was to first establish the presence of the phenomenon before (and alongside) identifying the factors that influence it. Turri et al. (2015) employed a particular strategy and methodology for studying knowledge attributions in their Experiment 1 and made conclusions based on the results. The purpose of Hall et al.’s replication was to use similar strategies to determine whether Turri et al.’s findings were robust. The replication results demonstrate that prior to studying mechanisms, establishing agreement on what researchers are trying to understand is necessary. Examining “specific forms of luck” and theorizing about their importance, as Buckwalter and Friedman recommended, would have been premature. When results do not conform to hypotheses, researchers tend to avoid accepting the conclusion that their theory may not be correct (i.e., refuted; Popper, 1959). Instead, not infrequently, they embark on subtle searches for subgroups or moderators for which their theory may hold. Hall et al. found differences and highlighted methodological issues, which, had they not been identified, could have hindered progress on theoretical development.

### Balancing Replication Goals

Buckwalter and Friedman (2024) suggested that Hall et al. (2024) did not appreciate the tension between two replication goals: that of assessing reliability and that of understanding phenomena. Hall et al. set out to do a simple, direct replication study with both pedagogical and scientific goals. Because of the partnership between the Psychological Science Accelerator (Moshontz et al., 2018) and the Collaborative Replications and Education Project (Grahe et al., 2024; Wagge et al., 2019), the research project increased in complexity and scope over the course of the Stage 1 Registered Report review process. As a consequence, additional piloted vignettes were incorporated, measures were added and changed, alternative response options were included, extension hypotheses were proposed by coauthors, and plans to assess moderating variables were formulated. The conceptual replication Hall et al. conducted emerged out of precisely those tensions noted by Buckwalter and Friedman and the additional consideration of pedagogical value. We hope that the findings of Hall et al. will contribute to the incremental understanding of epistemic intuitions and, in turn, inspire future inquiry along these lines, that is, examining subtypes, mechanisms, and moderators of so-called Gettier intuitions.

### Transparency

*Action Editor:* David A. Sbarra

*Editor:* David A. Sbarra

*Author Contributions*

**Kathleen Schmidt:** Conceptualization; Project administration; Writing – original draft; Writing – review & editing.

**Gerald J. Haeffel:** Conceptualization; Writing – original draft; Writing – review & editing.

**Neil Levy:** Conceptualization; Writing – review & editing.

**David Moreau:** Conceptualization; Writing – review & editing.

**Sean T. H. Lee:** Conceptualization; Writing – review & editing.

**Erin M. Buchanan:** Conceptualization; Writing – review & editing.

**Anthony J. Krafnick:** Conceptualization; Writing – review & editing.

**Martin Voracek:** Conceptualization; Writing – review & editing.

**Gerit Pfuhl:** Conceptualization; Writing – review & editing.

**Krystian Barzykowski:** Conceptualization; Writing – review & editing.

**Marta Kowal:** Conceptualization; Writing – review & editing.

**Jordan Wagge:** Conceptualization; Project administration; Writing – original draft; Writing – review & editing.











### 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.

### Funding

K. Schmidt was supported by a grant from the John Templeton Foundation (61825). N. Levy was supported by a grant from the John Templeton Foundation (62631). K. Barzykowski was supported by a grant from the National Science Centre, Poland (2019/35/B/HS6/00528). J. Wagge was supported by a grant from the National Science Foundation (2141930).

### ORCID iDs

Kathleen Schmidt  <https://orcid.org/0000-0002-9946-5953>  
 Gerald J. Haefel  <https://orcid.org/0000-0002-4029-1493>  
 Neil Levy  <https://orcid.org/0000-0002-5679-1986>  
 David Moreau  <https://orcid.org/0000-0002-1957-1941>  
 Sean T. H. Lee  <https://orcid.org/0000-0002-4059-9738>  
 Erin M. Buchanan  <https://orcid.org/0000-0002-9689-4189>  
 Anthony J. Krafnick  <https://orcid.org/0000-0002-1692-0413>  
 Martin Voracek  <https://orcid.org/0000-0001-6109-6155>  
 Marta Kowal  <https://orcid.org/0000-0001-9050-1471>  
 Jordan Wagge  <https://orcid.org/0000-0002-5105-2084>

### Note

1. In this analysis, luck attributions did not moderate the Gettier-intuition effect.

### References

- Buckwalter, W., & Friedman, O. (2024). Robust evidence for knowledge attribution and luck: A comment on Hall et al. (2024). *Advances in Methods and Practices in Psychological Science*, 7(4). <https://doi.org/10.1177/25152459241268220>
- Colaço, D., Buckwalter, W., Stich, S., & Machery, E. (2014). Epistemic intuitions in fake-barn thought experiments. *Episteme*, 11(2), 199–212. <https://doi.org/10.1017/epi.2014.7>
- Frick, R. W. (1996). The appropriate use of null hypothesis testing. *Psychological Methods*, 1(4), 379–390. <https://doi.org/10.1037/1082-989X.1.4.379>
- Grahe, J. E., Brandt, M. J., Wagge, J. R., Legate, N., Wiggins, B. J., Christopherson, C. D., Weisberg, Y., Corker, K. S., Chartier, C. R., Fallon, M., Hildebrandt, L., Hurst, M. A., Lazarevic, L., Levitan, C., McFall, J., McLaughlin, H., Pazda, A., IJzerman, H., Nosek, B. A., . . . France, H. (2024). *Collaborative Replications and Education Project (CREP)*. OSF. <https://doi.org/10.17605/OSF.IO/WFC6U>
- Hall, B. F., Schmidt, K., Wagge, J. R., Lewis, S. C., Weissgerber, S. C., Kiunke, F., Pfuhl, G., Stieger, S., Tran, U. S., Barzykowski, K., Bogatyreva, N., Kowal, M., Massar, K., Pernerstorfer, F., Sorokowski, P., Voracek, M., Chartier, C. R., Brandt, M. J., Grahe, J. E., . . . Buchanan, E. M. (2024). Registered Replication Report: A large multilab cross-cultural conceptual replication of Turri et al. (2015). *Advances in Methods and Practices in Psychological Science*, 7(4). <https://doi.org/10.1177/25152459241267902>
- Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., Grahe, J. E., McCarthy, R. J., Musser, E. D., Antfolk, J., Castille, C. M., Evans, T. R., Fiedler, S., Flake, J. K., Forero, D. A., Janssen, S. M. J., Keene, J. R., Protzko, J., Aczel, B., . . . Chartier, C. R. (2018). The Psychological Science Accelerator: Advancing psychology through a distributed collaborative network. *Advances in Methods and Practices in Psychological Science*, 1(4), 501–515. <https://doi.org/10.1177/2515245918797607>
- Nagel, J., San Juan, V., & Mar, R. A. (2013). Lay denial of knowledge for justified true beliefs. *Cognition*, 12(3), 652–661. <https://doi.org/10.1016/j.cognition.2013.02.008>
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences, USA*, 115(11), 2600–2606. <https://doi.org/10.1073/pnas.1708274114>
- Popper, K. R. (1959). *The logic of scientific discovery*. Basic Books.
- Turri, J., Buckwalter, W., & Blouw, P. (2015). Knowledge and luck. *Psychonomic Bulletin & Review*, 22(2), 378–390. <https://doi.org/10.3758/s13423-014-0683-5>
- Wagge, J. R., Brandt, M. J., Lazarevic, L. B., Legate, N., Christopherson, C., Wiggins, B., & Grahe, J. E. (2019). Publishing research with undergraduate students via replication work: The Collaborative Replications and Education Project. *Frontiers in Psychology*, 10, Article 439232. <https://doi.org/10.3389/fpsyg.2019.00247>
- Weinberg, J. M., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics*, 29(1), 429–460. <https://doi.org/10.5840/philtopics2001291/217>
- Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in Psychology*, 7, Article 1832. <https://doi.org/10.3389/fpsyg.2016.01832>