

Conclusion

We believe that our explanation of the similarities between the TOST procedure and the SGPV provides context to interpret the contribution of second generation p -values to the statistical toolbox. The novelty of the SGPV can be limited when confidence intervals are asymmetrical or wider than the equivalence range. There are strong similarities with p -values from the TOST procedure, and in all situations where the statistics yield different results, the behavior of the p -value from the TOST procedure is more consistent. We hope this overview of the relationship between the SGPV and equivalence tests will help researchers to make an informed decision about which statistical approach provides the best answer to their question. Our comparisons show that when proposing alternatives to null-hypothesis tests, it is important to compare new proposals to already existing procedures. We believe equivalence tests achieve the goals of the second generation p -value while allowing users to easily control error rates, and while yielding more consistent statistical outcomes.

Authors Note

All code associated with this article, including the reproducible manuscript, is available from https://github.com/Lakens/TOST_vs_SGPV and <https://osf.io/8crkg/>. The preprint can be found at <https://psyarxiv.com/7k6ay/>.

Correspondence concerning this article should be addressed to Daniël Lakens, Den Dolech 1, IPO 1.33, 5600 MB, Eindhoven, The Netherlands. E-mail: D.Lakens@tue.nl

Open Science Practices



This article earned the Open Materials badge for making the materials openly available. It has been verified that the analysis reproduced the results presented in the article. The entire editorial process, including the open reviews, are published in the online supplement.

Conflict of Interest and Funding

No conflict of interest and no external funding. This work was supported by the Netherlands Organization for Scientific Research (NWO) VIDI grant 452-17-013.

Author Contributions

DL conceptualized the idea, both authors wrote and revised this manuscript.

References

- American Psychological Association (Ed.). (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.
- Blume, J. D., D'Agostino McGowan, L., Dupont, W. D., & Greevy, R. A. (2018). Second-generation p -values: Improved rigor, reproducibility, & transparency in statistical analyses. *PLOS ONE*, *13*(3), e0188299. doi:[10.1371/journal.pone.0188299](https://doi.org/10.1371/journal.pone.0188299)
- Goertzen, J. R., & Cribbie, R. A. (2010). Detecting a lack of association: An equivalence testing approach. *British Journal of Mathematical and Statistical Psychology*, *63*(3), 527–537. doi:[10.1348/000711009X475853](https://doi.org/10.1348/000711009X475853)
- Hauck, D. W. W., & Anderson, S. (1984). A new statistical procedure for testing equivalence in two-group comparative bioavailability trials. *Journal of Pharmacokinetics and Biopharmaceutics*, *12*(1), 83–91. doi:[10.1007/BF01063612](https://doi.org/10.1007/BF01063612)
- Kruschke, J. K. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation. *Advances in Methods and Practices in Psychological Science*, *1*(2), 270–280. doi:[10.1177/2515245918771304](https://doi.org/10.1177/2515245918771304)
- Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. doi:[10.1177/1948550617697177](https://doi.org/10.1177/1948550617697177)
- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. doi:[10.1177/2515245918770963](https://doi.org/10.1177/2515245918770963)
- Meyners, M. (2012). Equivalence tests A review. *Food Quality and Preference*, *26*(2), 231–245. doi:[10.1016/j.foodqual.2012.05.003](https://doi.org/10.1016/j.foodqual.2012.05.003)
- Murphy, K. R., Myers, B., & Wolach, A. H. (2014). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (Fourth edition.). New York: Routledge, Taylor & Francis Group.
- Quertemont, E. (2011). How to Statistically Show the Absence of an Effect. *Psychologica Belgica*, *51*(2), 109–127. doi:[10.5334/pb-51-2-109](https://doi.org/10.5334/pb-51-2-109)
- Rogers, J. L., Howard, K. I., & Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin*, *113*(3), 553–565. doi:[http://dx.doi.org/10.1037/0033-2909.113.3.553](https://doi.org/http://dx.doi.org/10.1037/0033-2909.113.3.553)