A systematic review of Null Hypothesis Significance Testing,

sample sizes and statistical power in research using

the Implicit Relational Assessment Procedure

*Supplementary Materials*

# Systematic review full text exclusions with reasons

Publications by Baker et al. (2015), Baker et al. (2017), Smith et al. (2022) and Szarko et al. (2022) each reported employing an IRAP but did not report the procedural features in sufficient detail to determine whether it was a standard IRAP or a variant. Personal correspondence with the authors revealed these studies employed a non-standard IRAP variant. Following the inclusion criteria, these studies were therefore excluded.

Inspection of the full text of Perez et al. (2020) demonstrated that those studies did not employ an IRAP (or IRAP variant).

Inspection of the full texts of four articles by Harte and colleagues (Harte, Barnes-Holmes, et al., 2021; Harte, Barnes‐Holmes, et al., 2021; Harte et al., 2018, 2020)(2018, 2020, 2021a, 2021b) demonstrated that those studies employed a Training IRAP (T-IRAP) rather than a standard IRAP. Following the inclusion criteria, these studies were therefore excluded.

# Tables

**Table S1.** Number of sample sizes in the Social and Personality psychology dataset by journal

| Journal | *N* sample sizes |
| --- | --- |
| European Journal of Personality | 113 |
| European Journal of Social Psychology | 306 |
| Journal of European Social Psychology | 589 |
| Journal of Personality | 189 |
| Journal of Personality and Social Psychology | 631 |
| Journal Research in Personality | 157 |
| Psychological Science (articles coded as relevant to Social Psychology) | 269 |
| Personality and Social and Psychology Bulletin | 527 |
| Social Psychological and Personality Science | 266 |

**Table S2.** Median sample size per year in IRAP studies

|  | Year | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Calculated from | ’06 | ’07 | ’08 | ’09 | ’10 | ’11 | ’12 | ’13 | ’14 | ’15 | ’16 | ’17 | ’18 | ’19 | ’20 | ’21 | ’22 |
| Participants per study | 14 (4) | 36 (1) | 12 (1) | 29 (10) | 19 (13) | 64 (5) | 28 (6) | 41 (7) | 32 (9) | 41 (16) | 44 (28) | 32 (17) | 44 (15) | 37 (19) | 50 (16) | 41 (11) | 64 (7) |
| Participants per group in studies with between-subjects comparisons | 15 (1) | 12 (1) | 6 (1) | 13 (9) | 11 (6) | 17 (3) | 15 (3) | 21 (5) | 14 (8) | 18 (12) | 20 (22) | 16 (11) | 22 (9) | 19 (10) | 24 (12) | 19 (7) | 30 (5) |

This table corresponds with Figure 2. The number of studies each median is calculated from is listed in brackets.

**Table S3.** The statistical power to detect the average published effect size (Cohen’s *d* = 0.408 or Pearson’s *r* = 0.20) implied by median sample size per group per year in IRAP studies

| Year | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ’06 | ’07 | ’08 | ’09 | ’10 | ’11 | ’12 | ’13 | ’14 | ’15 | ’16 | ’17 | ’18 | ’19 | ’20 | ’21 | ’22 |
| 19 | 16 | 10 | 17 | 15 | 21 | 19 | 25 | 18 | 22 | 24 | 20 | 26 | 23 | 28 | 23 | 34 |

This table corresponds with Figure 3.

**Table S4.** Median sample size per year in IRAP studies compared to Social and Personality Psychology studies

|  | Year | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ’06 | ’07 | ’08 | ’09 | ’10 | ’11 | ’12 | ’13 | ’14 | ’15 | ’16 | ’17 | ’18 | ’19 | ’20 | ’21 | ’22 |
| Social and Personality  research | - | - | - | - | - | 103 (314) | 98 (304) | 130 (316) | 124 (297) | 153 (390) | 156 (327) | 221 (344) | 227 (402) | 251 (353) | - | - | - |
| IRAP research | 44 (1) | 36 (1) | 12 (1) | 26 (9) | 22 (6) | 64 (3) | 30 (3) | 41 (5) | 36 (8) | 43 (12) | 48 (22) | 32 (11) | 49 (9) | 46 (10) | 53 (12) | 37 (7) | 64 (5) |

This table corresponds with Figure 4. The number of studies each median is calculated from is listed in brackets.

**Table S5.** The statistical power to detect the average published effect size (Cohen’s *d* = 0.408 or Pearson’s *r* = 0.20) implied by median sample size per study per year in IRAP studies compared to Social and Personality Psychology studies

|  | Year | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ’06 | ’07 | ’08 | ’09 | ’10 | ’11 | ’12 | ’13 | ’14 | ’15 | ’16 | ’17 | ’18 | ’19 | ’20 | ’21 | ’22 |
| Social and Personality  research | - | - | - | - | - | 53 | 51 | 63 | 61 | 70 | 71 | 85 | 86 | 89 | - | - | - |
| IRAP research | 26 | 22 | 10 | 17 | 15 | 36 | 19 | 24 | 22 | 25 | 28 | 20 | 28 | 27 | 30 | 22 | 36 |

This table corresponds with Figure 5.

References

Baker, T. K., Schwenk, T., Piasecki, M., Smith, G. S., Reimer, D., Jacobs, N., Shonkwiler, G., Hagen, J., & Houmanfar, R. A. (2015). Cultural Change in a Medical School: A Data-Driven Management of Entropy. *Journal of Organizational Behavior Management*, *35*(1–2), 95–122. https://doi.org/10.1080/01608061.2015.1035826

Baker, T. K., Smith, G. S., Jacobs, N. N., Houmanfar, R., Tolles, R., Kuhls, D., & Piasecki, M. (2017). A deeper look at implicit weight bias in medical students. *Advances in Health Sciences Education*, *22*(4), 889–900. https://doi.org/10.1007/s10459-016-9718-1

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2018). The impact of high versus low levels of derivation for mutually and combinatorially entailed relations on persistent rule-following. *Behavioural Processes*, *157*, 36–46. APA PsycInfo. https://doi.org/10.1016/j.beproc.2018.08.005

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2021). Exploring the impact of coherence (through the presence versus absence of feedback) and levels of derivation on persistent rule-following. *Learning & Behavior*, *49*(2), 222–239. APA PsycInfo. https://doi.org/10.3758/s13420-020-00438-1

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., McEnteggart, C., Gys, J., & Hasler, C. (2020). Exploring the potential impact of relational coherence on persistent rule-following: The first study. *Learning & Behavior*, *48*(3), 373–391. APA PsycInfo. https://doi.org/10.3758/s13420-019-00399-0

Harte, C., Barnes‐Holmes, D., Moreira, M., de Almeida, J. H., Passarelli, D., & de Rose, J. C. (2021). Exploring a training IRAP as a single participant context for analyzing reversed derived relations and persistent rule‐following. *Journal of the Experimental Analysis of Behavior*, *115*(2), 460–480. APA PsycInfo. https://doi.org/10.1002/jeab.671

Perez, W. F., de Almeida, J. H., Soares, L. C. C. S., Wang, T. F. L., de Morais, T. E. D. G., Mascarenhas, A. V., & de Rose, J. C. (2020). Fearful Faces and the Derived Transfer of Aversive Functions. *The Psychological Record*. https://doi.org/10.1007/s40732-020-00390-6

Smith, G. S., Houmanfar, R. A., Jacobs, N. N., Froehlich, M., Szarko, A. J., Smith, B. M., Kemmelmeier, M., Baker, T. K., Piasecki, M., & Schwenk, T. L. (2022). Assessment of medical student burnout: Toward an implicit measure to address current issues. *Advances in Health Sciences Education*, *27*(2), 375–386. https://doi.org/10.1007/s10459-021-10089-0

Szarko, A. J., Houmanfar, R. A., Smith, G. S., Jacobs, N. N., Smith, B. M., Assemi, K., Piasecki, M., & Baker, T. K. (2022). Impact of Acceptance and Commitment Training on psychological flexibility and burnout in medical education. *Journal of Contextual Behavioral Science*, *23*, 190–199. APA PsycInfo. https://doi.org/10.1016/j.jcbs.2022.02.004