

Lower your expectations: Effect sizes in psychology

Live google doc [here](#)

Proposal for third year paper here [here](#)

This is a working document on a potential research project about the true distribution of effect size in psychology and its determinants.

There are three possibilities:

1. There is a potential empirical contribution, a meta-meta analysis. We would analyze published meta-analyses and examine the distribution of effect sizes and their determinants.
2. There is a potential contribution of a review/essay/perspective paper. We would synthesize the literature on effect size in psychology and discuss the implications of the findings.
3. We will conclude that there is no value in pursuing this project.

For now, we will focus on the **empirical** option, which is our working hypothesis.

Motivation

There is a lot that has been said about effect sizes in psychology. (1) There are a bunch of essays that argue that effect sizes are smaller than Cohen's traditional cutoffs, and are still important (e.g.,^{1 2}). (2) There are some meta-analyses for particular subdisciplines, or areas of research (e.g., cognitive neuroscience³, nudges⁴, applied psychology⁵). (3) Close to what we want to achieve, there is an estimation of the magnitude of the file drawer effect⁶. (4) The many labs replication is also relevant, in that it showed that across 100 published experiments, their average published effect size was $r = .40$, whereas the well-powered replication effect sizes were roughly half of this⁷. (5) There are also empirical studies demonstrating that even with plenty of data and sophisticated methods life outcomes are really hard to predict⁸. (A similar in principle simulation study⁹).

Despite all this, there is no definitive answer for a scientist planning a study and trying to answer the question of what effect size should I expect? How should I plan to power my study? Our vision is that our study would be able to model effect sizes such that I could know that the median effect size in psychology is, say, $r = .15$. If I know that we are dealing with a social psychology effect, where the predictor and the outcome do not share common method variance (e.g., open ended text correlating with behavior), and the predictor and outcome are separated by 4 years, then, I can expect my effect size to shrink to $r = .06$.

Aside from the contribution towards sample size planning and power calculations, we also hope to show how miscalibrated researchers are in terms of what effect sizes can be expected, and how we should evaluate them when reading, reviewing, and conducting research.

Research questions

- What is the distribution of effect sizes in psychology?
- How does this distribution change for different types of studies?
 - Published vs. unpublished
 - Lab studies vs. field studies
 - Experimental vs. correlational studies
 - Social vs. cognitive
 - Within person vs. between person
 - Concurrent vs. time-lagged
 - Common method vs multi-method
- How miscalibrated are scientists to true effect sizes. For what kinds of research are they more likely to overestimate effect sizes? What causes this miscalibration?

Approach

- We will use meta-analyses as our unit of analysis. We will use the meta-analytic effect size as our dependent variable.
- We could download all meta-analyses published in Psychological Bulletin and code them for the above variables.

Open questions

- Is this a tractable problem?
- If it were possible, would it be a useful contribution?

- We expect to have a missing data problem with some of the moderators of effect size. How big of a problem would this be?
- Given what already is out there, is it true that this is a case of “go big or go home”?
- Would it be useful to also use metaBUS for this?

Extra notes from Literature review

Dalton et al¹⁰ suggests that there is **no** publication bias or file drawer inflation bias in personnel psychology. - I don't believe that!

Notes from Meetings

- Dolores thinks that just using Psych Bulletin data is not gonna fly in Psych Bulletin. It could work for PNAS or some other journal she said.
 - An interesting point from Dolores was to use the estimated distribution to try to uncover insights from psychology. To interesting examples: (1) Are effects larger for younger samples? Are kids more malleable? (2) Are effects larger the closer you get to behavior.
1. Götz, F. M., Gosling, S. D., & Rentfrow, P. J. (2022). *Small Effects: The Indispensable Foundation for a Cumulative Psychological Science*.
 2. Funder, D. C., & Ozer, D. J. (2019). Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2), 156–168. <https://doi.org/10.1177/2515245919847202>
 3. Szucs, D., & Ioannidis, J. P. A. (2017). Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLOS Biology*, 15(3), e2000797. <https://doi.org/10.1371/journal.pbio.2000797>
 4. DellaVigna, S., & Linos, E. (2022). RCTs to Scale: Comprehensive Evidence From Two Nudge Units. *Econometrica*, 90(1), 81–116. <https://doi.org/10.3982/ECTA18709>
 5. Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, 100(2), 431–449. <https://doi.org/10.1037/a0038047>
 6. Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the difference between published and unpublished effect sizes: A meta-review. *Review of Educational Research*, 86(1), 207–236. <https://doi.org/10.3102/0034654315582067>
 7. Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. <https://doi.org/10.1126/science.aac4716>
 8. Salganik, M. J., Lundberg, I., Kindel, A. T., Ahearn, C. E., Al-Ghoneim, K., Almaatouq, A., Altschul, D. M., Brand, J. E., Carnegie, N. B., Compton, R. J., Datta, D., Davidson, T., Filippova, A., Gilroy, C., Goode, B. J., Jahani, E., Kashyap, R., Kirchner, A., McKay, S., ... McLanahan, S. (2020). Measuring the predictability of life outcomes with a scientific

- mass collaboration. *Proceedings of the National Academy of Sciences*, 117(15), 8398–8403. <https://doi.org/10.1073/pnas.1915006117>
9. Ahadi, S., & Diener, E. (1989). Multiple determinants and effect size. *Journal of Personality and Social Psychology*, 56(3), 398–406. <https://doi.org/10.1037/0022-3514.56.3.398>
 10. Dalton, D. R., Aguinis, H., Dalton, C. M., Bosco, F. A., & Pierce, C. A. (2012). Revisiting the file drawer problem in meta-analysis: An assessment of published and nonpublished correlation matrices: Personnel Psychology. *Personnel Psychology*, 65(2), 221–249. <https://doi.org/10.1111/j.1744-6570.2012.01243.x>