

No reason to expect large and consistent effects of nudge interventions

Szaszi, B.*, Higney, A., Charlton, A., Gelman, A., Ziano, I., Aczel, B., Goldstein, D. G., Yeager, D. S., Tipton, E.

As policy-makers are increasingly interested in implementing nudge-type interventions, it is essential that we understand under what conditions they can improve policy-relevant outcomes to make the best possible use of public resources. For that reason, the recently published meta-analysis by Mertens et al.¹ of the choice architecture literature is laudable.

Our reading of the data and analyses, however, is quite different from Mertens et al.'s: nudge interventions may work, under certain conditions, but their effectiveness can vary a great degree and the conditions under which they work are barely identified in the literature². For example, the authors assume that the nudge literature is impacted by publication bias; that is, larger positive, and statistically significant comparisons are more likely to be reported. After adjusting for hypothesized severe to moderate degree of publication bias, their adjusted estimated average effect of nudges is between $d = 0.03$ (severe) to $d = 0.31$ (moderate). Our additional analysis on the same database applying three different bias-correcting methods, compared to the non-adjusted estimate, also led to much smaller effect sizes (Andrews-Kasy, $d = -0.01$, $se = 0.01$; WAAP, $d = 0.06$, $se = 0.03$; Trim and fill, $d = 0.08$, $se = 0.03$)^{see also 3}. Furthermore, the authors estimate that even after adjusting for publication bias, the effects of nudge interventions vary considerably across studies. For example, assuming a severe degree of publication bias, 95% of these studies' effects would be +/- 1.14 around the average $d = 0.03$ effect showing large variability, with much of this variability possibly arising from variability in publication bias itself.

Nevertheless, Mertens et al. focus their message on the average effect size estimated without adjusting for publication bias, concluding that “our results show that choice architecture interventions promote behavior change with a small to medium effect size of Cohen’s $d = 0.45$ ”(p. 1). We argue that this effect size is implausibly large which could be misleading and further strengthen researchers’ and practitioners’ overoptimistic expectations^{3,4} about the impact of nudges. Furthermore, the authors focus their conclusions on this average value and on subgroups, leaving aside the large degree of unexplained heterogeneity⁵ in apparent effects across published studies. For example, despite the analyses above being consistent with a large proportion of studies having near-zero underlying effects, they conclude that nudges work “across a wide range of behavioral domains, population segments, and geographical locations” (p. 7).

Thankfully, it is because Mertens et al. conducted these analyses and shared their data that we were able to notice these contradictions between findings and conclusions. We argue that as a scientific field, instead of focusing on average effects, we need to understand when and where some nudges have huge positive effects and why others are not able to repeat those

successes^{2,4,5}. Until then, with a few exceptions (e.g., defaults⁶), we see no reason to expect large and consistent effects when designing nudge experiments or running interventions.

References

1. Mertens S, Herberz M, Hahnel UJ, Brosch T. The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proc Natl Acad Sci*. 2022;119(1).
2. Szaszi B, Palinkas A, Palfi B, Szollosi A, Aczel B. A Systematic Scoping Review of the Choice Architecture Movement: Toward Understanding When and Why Nudges Work. *J Behav Decis Mak*. 2018;31(3):355-366. doi:10.1002/bdm.2035
3. DellaVigna S, Linos E. *Rcts to Scale: Comprehensive Evidence from Two Nudge Units*. National Bureau of Economic Research; 2020.
4. Milkman KL, Gromet D, Ho H, et al. Megastudies improve the impact of applied behavioural science. *Nature*. 2021;600(7889):478-483.
5. Bryan CJ, Tipton E, Yeager DS. Behavioural science is unlikely to change the world without a heterogeneity revolution. *Nat Hum Behav*. 2021;5(8):980-989.
6. Jachimowicz JM, Duncan S, Weber EU, Johnson EJ. When and why defaults influence decisions: A meta-analysis of default effects. *Behav Public Policy*. 2019;3(2):159-186.

Open data

<https://osf.io/bntm6/>