



**Brian Nosek** @BrianNosek · 19h

Many Labs 2: 28 findings, 60+ samples, ~7000 participants each study, 186 authors, 36 nations.

Successfully replicated 14 of 28 [psyarxiv.com/9654g](https://psyarxiv.com/9654g)



**Brian Nosek** @BrianNosek · 19h

ML2 minimized boring reasons for failure. First, using original materials & Registered Reports [cos.io/rr](https://cos.io/rr) all 28 replications met expert reviewed quality control standards. Failure to replicate not easily dismissed as replication incompetence. [psyarxiv.com/9654g](https://psyarxiv.com/9654g)



**Brian Nosek** @BrianNosek · 19h

Second, the total ML2 replication median sample size ( $n = 7157$ ) was 64x original median sample size ( $n = 112$ ). If there was an effect to detect, even a much smaller one, we would detect it. Ultimate estimates have very high precision. [psyarxiv.com/9654g](https://psyarxiv.com/9654g)



**Brian Nosek** @BrianNosek · 19h

Third, each original finding was replicated in >60 samples with labs from 36 nations and territories taking part. If the effect was easy to detect in some samples and not others, ML2 would find evidence for that. [psyarxiv.com/9654g](https://psyarxiv.com/9654g)



**Brian Nosek** @BrianNosek · 19h

Fourth, some original authors offered a priori hypotheses of moderating influences based on sample, task order, or other design features. Identifying them a priori provided an opportunity to test them with confidence.



**Brian Nosek** @BrianNosek · 19h

The top line results of 50% overall replication success rate and effect sizes less than 1/2 of original studies are consistent with prior replication studies. [psyarxiv.com/9654g](https://psyarxiv.com/9654g)

△ Original Effect Size

Cohen's q

-3 -2 -1 0 1 2 3

