nature human behaviour

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41562-019-0787-z

In the format provided by the authors and unedited.

Comparing meta-analyses and preregistered multiple-laboratory replication projects

Amanda Kvarven[1,3], Eirik Strømland[1,3] and Magnus Johannesson[2]*

[1]Department of Economics, University of Bergen, Bergen, Norway. [2]Department of Economics, Stockholm School of Economics, Stockholm, Sweden. [3]These authors contributed equally: Amanda Kvarven, Eirik Strømland. *e-mail: magnus.johannesson@hhs.se

Supplementary Information for

Comparing Meta-Analyses and Pre-Registered Multiple Labs Replication Projects

Amanda Kvarven[1,†], Eirik Strømland[2,†], Magnus Johannesson[3*]

[1] Department of Economics, University of Bergen, P.O. 7802, 5020 Bergen, Norway

[2] Department of Economics, University of Bergen, P.O. 7802, 5020 Bergen, Norway

[3] Department of Economics, Stockholm School of Economics, Box 6501, SE-113 83 Stockholm, Sweden

[*] Corresponding author: Magnus Johannesson (magnus.johannesson@hhs.se).

[†] These authors contributed equally to this work.

The Supplementary Information contain:

Supplementary Tables 1-16

Supplementary References

Supplementary Table 1. List of all studies that were assessed for eligibility for inclusion in our analysis sample.

Original	Replication	Meta-analysis	Note on meta-analysis	Included
Critcher & Gilovich (2008)[1]	Klein et al. (2018)[2]	Henriksson (2015)[3]	Main estimate included	Yes
Graham, Haidt & Nosek (2009)[4]	Klein et al. (2018)[2]	Kivikangas, Lönnqvist & Ravaja (2017)[5]	Authority estimate included, others used in robustness checks	Yes
Hauser et al. (2007)[6]	Klein et al. (2018)[2]	Feltz & May (2017)[7]	Estimate from categorical data used	Yes
Husnu & Crisp (2010)[8]	Klein et al. (2014)[9]	Miles & Crisp (2014)[10]	Sub-analysis (intentions) included in analysis	Yes
Jostmann, Lakens & Schubert (2009)[11]	Ebersole (2016)[12]	Rabelo et al. (2015)[13]	Main estimate included	Yes
Mazar et al. (2008)[14]	Verschuere et al. (2018)[5,16]	Belle & Cantarelli (2017)[17]	Estimate from appendix included	Yes
Monin & Miller (2001)[18]	Ebersole (2016)[12]	Blanken, van de Ven & Zeelenberg (2015)[19]	Main estimate included	Yes
Oppenheimer, Meyvis & Davidenko (2009)[20]	Klein et al. (2014)[9]	Roth, Robbert & Straus (2014)[21]	Sub-analysis (sunk costs) included in analysis	Yes
Rand, Greene & Nowak (2012)[22]	Bouwmeester et al. (2017)[23]	Rand (2016)[24]	ITT estimate included in analysis	Yes
Schooler & Engstler-Schooler (1990)[25]	Alogna et al. (2014)[26]	Meissner & Brigham (2001)[27]	Main estimate included	Yes
Schwarz, Strack & Mai (1991)[28]	Klein et al. (2018)[2]	Schimmack & Oishi (2005)[29]	Random effects estimate included	Yes
Sripada, Kessler & Jonides (2014)[30]	Hagger et al. (2016)[31]	Hagger et al. (2010)[32]	Overall ego-depletion effect used	Yes
Sripada, Kessler & Jonides (2014)[30]	Hagger et al. (2016)[31]	Carter et al. (2015)[33]	Included in robustness check in replacement of Hagger et al. (2010). Main random effects estimate used	In robustness check
Srull & Wyer (1979)[34]	McCarthy et al. (2018)[16]	DeCoster & Claypool (2004)[35]	Main estimate included, sub-analysis does not report all relevant info	Yes
Strack, Martin & Stepper (1988)[36]	Wagenmakers et al. (2016)[37]	Coles et al. (2019)[38]	Main estimate included, Wagenmakers et al. (2016) was a large part of relevant sub-analysis	Yes
Tversky & Kahneman (1981)[39]	Klein et al. (2014)[9]	Kühberger (1998)[40]	Sub-analysis (Asian disease) included in analysis	Yes
Anderson et al. (2012)[41]	Klein et al. (2018)[2]	Pinquart & Sörensen (2000)[42]	Excluded: No match	No
Boroditsky (2000)[43]	Ebersole (2016)[12]	Stickles & Lewis (2018)[44]	Excluded: Meta-analysis does not look at the difference between treatment and control, only average in control group	No
Cacioppo, Petty & Morris (1983)[45]	Ebersole (2016)[12]	Von Stumm & Ackerman (2013)[46]	Excluded: No match	No
Cacioppo, Petty & Morris (1983)[45]	Ebersole (2016)[12]	Stiff (1986)[47]	Excluded: No match	No
Caruso, Vohs, Baxter & Waytz (2013)[48]	Klein et al. (2014)[9]	Vadillo, Hardwicke, & Shanks (2016)[49]	Excluded: Not a proper meta-analysis	No
Caruso, Vohs, Baxter & Waytz (2013)[48]	Klein et al. (2014)[9]	Schuler & Wänke (2016)[50]	Excluded: Not a proper meta-analysis	No
Finkel et al. (2002)[51]	Cheung et al. (2016)[52]	Fehr et al. (2010)[53]	Excluded: No match	No
Galinsky et al. (2006)[54]	Ebersole (2016)[12]	Hall et al. (2015)[55]	Excluded: No match	No
Graham, Haidt & Nosek (2009)[4]	Klein et al. (2018)[2]	Skitka, Morgan & Wisneski (2015)[56]	Excluded: No match	No
Inbar et al. (2009)[57]	Klein et al. (2018)[2]	Terrizzi, Shook & McDaniel (2013)[58]	Excluded: No match	No
Inbar et al. (2009)[57]	Klein et al. (2018)[2]	Kiss, Morrison & Morrison (2018)[59]	Excluded: No match, not same outcome variable	No
Jacowitz & Kahneman (1995)[60]	Klein et al. (2014)[9]	DeCoster & Claypool (2014)[35]	Excluded: No match	No
Jacowitz & Kahneman (1995)[60]	Klein et al. (2014)[9]	Orr & Guthrie (2006)[61]	Excluded: No match	No
Nosek, Banaji & Greenwald (2002)[62]	Klein et al. (2014)[9]	Greenwald et al. (2009)[63]	Excluded: No match	No
Nosek, Banaji & Greenwald (2002)[62]	Klein et al. (2014)[9]	Hofmann et al. (2005)[64]	Excluded: No match	No
Nosek, Banaji & Greenwald (2002)[62]	Klein et al. (2014)[9]	Lindberg et al. (2010)[65]	Excluded: No match	No
Nosek, Banaji & Greenwald (2002)[62]	Klein et al. (2014)[9]	Reilly, Neumann & Andrews (2015)[66]	Excluded: No match	No
Ross, Greene & House (1977)[67]	Klein et al. (2018)[2]	Mullen (1985)[68]	Excluded: Data not available	No
Rugg (1941)[69]	Klein et al. (2014)[9]	Holleman (1999)[70]	Excluded: Data not available	No
Schooler & Engstler-Schooler (1990)[25]	Alogna et al. (2014)[26]	Meissner, Sporer & Susa (2008)[71]	Excluded: No match	No
Schwarz, Strack & Mai (1991)[28]	Klein et al. (2018)[2]	Heller, Watson & Ilies (2004)[72]	Excluded: No match	No
Zhong & Liljenquist (2006)[73]	Klein et al. (2018)[2]	Blanken, van de Ven & Zeelenberg (2015)[19]	Excluded: No match	No
Strack, Martin & Stepper (1988)[36]	Wagenmakers et al. (2016)[37]	Matsumoto (1987)[74]	Excluded: No match, not same outcome variable	No

Supplementary Table 2. Data for the original studies forming the basis for our data analysis. The effect sizes of the original studies have been defined as having a positive sign, irrespective of the sign used in the original study. The effect sizes are in Cohen's d units, unless otherwise noted in the Notes below the table.

ID	Reference	Sample size	Effect size	Standard error	95% CI	99.5% CI	Z-value	p-value	Effect size converted by us
1	Oppenheimer et al. (2009)[20]	N = 106	0.23	0.138	-0.040, 0.500	-0.157, 0.617	1.667	0.095	No
2	Tversky & Kahneman (1981)[39]	N = 152	1.13	0.122	0.891, 1.369	0.788, 1.472	9.262	<0.001	No
3	Husnu & Crisp (2010)[8]	N = 33	0.86	0.365	0.145, 1.575	-0.165, 1.885	2.356	0.018	No
4	Schwarz et al. (1991)[28]	N = 50	0.48	0.207	0.074, 0.886	-0.101, 1.061	2.319	0.020	No
5	Hauser et al. (2007)[6]	N = 2612, N = 2646	0.95	0.085	0.783, 1.117	0.711, 1.189	11.176	< 0.001	No
6	Critcher & Gilovich (2008)[1]	N = 207	0.30	0.143	0.020, 0.580	-0.101, 0.701	2.098	0.036	No
7	Graham, Haidt & Nosek (2009)[4]	N = 1548	0.52	0.059	0.404, 0.636	0.354, 0.686	8.814	<0.001	No
8	Jostmann et al. (2009)[11]	N = 51	0.597	0.293	0.023, 1.171	-0.225, 1.419	2.038	0.042	No
9	Monin & Miller (2001)[18]	N = 202	0.424	0.150	0.130, 0.718	0.003, 0.845	2.827	0.005	Yes
10	Schooler & Engstler-Schooler (1990)[25]	N = 88, N = 117	0.498	0.169	0.167, 0.829	0.024, 0.972	2.947	0.003	No
11	Sripada et al. (2014)[30]	N = 108	0.29	0.293	-0.284, 0.864	-0.532, 1.112	0.990	0.024	No
12	Rand et al. (2012)[22]	N = 680	0.22	0.138	-0.050, 0.490	-0.167, 0.607	1.594	0.111	Yes
13	Strack et al. (1988)[36]	N = 92	0.46	0.245	-0.020, 0.940	-0.228, 1.148	1.88	0.060	Yes
14	Srull & Wyer (1979)[34]	N = 43	3.01	N/A	N/A	N/A	N/A	N/A	No
15	Mazar et al. (2008)[14]	N = 229	0.48	0.196	0.096, 0.864	-0.070, 1.030	2.450	0.014	Yes

Notes: For Srull & Wyer (1979)[34] data on the standard deviation is unavailable, so the effect cannot be converted to Cohen's d as the raw effect needs to be divided by the standard deviation. For Schwarz et al. (1991)[28] the effect size is denoted in Cohen's q units, rather than Cohen's d units. For Hauser et al. (2007)[6] there are two relevant original estimates as both the replication and the meta-analysis looks at both the difference between scenario 1 and 2 and the difference between scenario 3 and 4. Hauser et al. (2007)[6] is therefore a pooled estimates based on these two differences. For Schooler & Engstler-Schooler (1990)[25] both the replication and the meta-analysis looks at study 1 and study 4. Schooler & Engstler-Schooler (1990)[25] is therefore a pooled estimate based on these two studies.

Supplementary Table 3. Data for the meta-analyses included in our analysis sample. The effect sizes are in Cohen's d units, unless otherwise noted in the Notes below the table. We also report the smallest detectable effect size with 80% power, denoted by "MDE (5%)" and "MDE (0.5%)" for the 15 meta-analyses (for both the 0.5% and the 5% significance levels).

ID	Original study	Sample size (K, N)	Effect size	Standard error	Z-value	p-value	95% CI	99.5% CI	Tau squared	MDE 5%	MDE 0.5%	Effect size converted by us
1	Oppenheimer et al. (2009)[20]	100, 15481	0.496	0.067	7.365	<0.001	0.365, 0.627	0.308, 0.684	0.368	0.188	0.245	No
2	Tversky & Kahneman (1981)[39]	80, 10789	0.57	0.02	27.929	<0.001	0.531, 0.609	0.514, 0.626	0.066	0.056	0.073	No
3	Husnu & Crisp (2010)[8]	32, 2076	0.459	0.069	6.664	<0.001	0.324, 0.594	0.265, 0.653	0.08	0.193	0.252	No
4	Schwarz et al. (1991)[28]	16, 1696	0.127	0.064	1.990	0.047	0.002, 0.252	-0.053, 0.307	0.156	0.179	0.234	Yes
5	Hauser et al. (2007)[6]	30, 24058	0.877	0.122	7.163	<0.001	0.638, 1.116	0.535, 1.219	0.54	0.342	0.445	Yes
6	Critcher & Gilovich (2008)[1]	21, 2165	0.31	0.071	4.340	<0.001	0.171, 0.449	0.111, 0.509	0.045	0.199	0.259	No
7	Graham, Haidt & Nosek (2009)[4]	49, 212521	0.676	0.048	14.073	<0.001	0.582, 0.770	0.541, 0.811	0.08	0.134	0.175	Yes
8	Jostmann et al. (2009)[11]	25, 1625	0.57	0.051	11.172	<0.001	0.470, 0.670	0.427, 0.713	0	0.143	0.186	No
9	Monin & Miller (2001)[18]	91, 7397	0.31	0.038	8.101	<0.001	0.236, 0.384	0.203, 0.417	0.06	0.106	0.139	No
10	Schooler & Engstler-Schooler (1990)[25]	29, 2018	0.236	0.046	5.168	<0.001	0.146, 0.326	0.107, 0.365	0.055	0.129	0.168	Yes
11	Sripada et al. (2014)[30]	198, 10782	0.62	0.026	24.304	<0.001	0.569, 0.671	0.548, 0.693	0.012	0.073	0.095	No
12	Rand et al. (2012)[22]	51, 15850	0.056	0.017	3.377	0.001	0.023, 0.089	0.008, 0.104	0.009	0.048	0.062	Yes
13	Strack et al. (1988)[36]	286, 23282	0.2	0.031	6.533	<0.001	0.139, 0.261	0.113, 0.287	0.11	0.087	0.113	Yes
14	Srull & Wyer (1979)[34]	45, 4794	0.3541	0.03	11.925	<0.001	0.295, 0.412	0.270, 0.438	0.035	0.084	0.110	No
15	Mazar et al. (2008)[14]	15, 1493	0.426	0.069	6.185	<0.001	0.291, 0.561	0.232, 0.620	0.014	0.193	0.252	No

Notes: As far as possible, the number of effect sizes (K) and the number of participants (N) is displayed for the exact estimate derived from the meta-analysis. For Hauser et al. (2007)[6] the general sample size is reported, not the sample size corresponding to the estimate picked for inclusion in our analysis sample. The effect size of the meta-analysis corresponding to the original study by Schwarz et al. (1991)[28] is denoted in Cohen's q units, rather than Cohen's d units. The effect size of this meta-analysis was converted to Cohen's q units by first converting the result of each individual study included in the meta-analysis to Cohen's q units and then estimating a random effects meta-analysis. The effect size of the meta-analyses corresponding to the original study by Critcher & Gilovich (2008)[1] and Mazar et al. (2008)[14] are denoted in Hedges' g units, rather than Cohen's d units.

Supplementary Table 4. Data for the replication studies included in our analysis sample. The effect sizes are in Cohen's d units, unless otherwise noted in the Notes below the table. We also report the smallest detectable effect size with 80% power, denoted by "MDE (5%)" and "MDE (0.5%)" for the 15 replication studies (for both the 0.5% and the 5% significance levels).

ID	Original study	Sample size (K, N)	Effect size	Standard error	Z-value	p-value	95% CI	99.5% CI	Effect size converted by us	MDE (5%)	MDE (0.5%)
1	Oppenheimer et al. (2009)[20]	36, 6330	0.27	0.036	7.560	<0.001	0.199, 0.341	0.169, 0.371	No	0.101	0.131
2	Tversky & Kahneman (1981)[39]	36, 6271	0.60	0.036	16.800	<0.001	0.529, 0.671	0.499, 0.701	No	0.101	0.131
3	Husnu & Crisp (2010)[8]	36, 6336	0.13	0.031	4.247	<0.001	0.069, 0.191	0.043, 0.217	No	0.087	0.113
4	Schwarz et al. (1991)[28]	61, 7460	-0.07	0.023	-3.030	0.002	-0.115, -0.025	-0.135, -0.005	No	0.064	0.084
5	Hauser et al. (2007)[6]	64, 6842 (Slate 1), 61, 7923 (Slate 2)	0.785	0.055	14.379	<0.001	0.677, 0.893	0.631, 0.939	No	0.154	0.201
6	Critcher & Gilovich (2008)[1]	64, 6826	0.04	0.026	1.568	0.117	-0.011, 0.091	-0.033, 0.113	No	0.073	0.095
7	Graham, Haidt & Nosek (2009)[4]	64, 6966	0.29	0.023	12.631	<0.001	0.245, 0.335	0.225, 0.355	No	0.064	0.084
8	Jostmann et al. (2009)[11]	20, 2285	0.03	0.043	0.692	0.489	-0.054, 0.114	-0.091, 0.151	No	0.120	0.157
9	Monin & Miller (2001)[18]	20, 3134	0.147	0.018	8.209	<0.001	0.112, 0.182	0.096, 0.198	Yes	0.050	0.066
10	Schooler & Engstler-Schooler (1990)[25]	31, 2603 22, 1535	0.076	0.014	5.476	<0.001	0.049, 0.103	0.037, 0.115	Yes	0.039	0.051
11	Sripada et al. (2014)[30]	23, 2141	0.04	0.056	0.713	0.476	-0.070, 0.150	-0.117, 0.197	No	0.157	0.204
12	Rand et al. (2012)[22]	21, 3596	-0.02	0.034	-0.581	0.561	-0.087, 0.047	-0.115, 0.075	Yes	0.095	0.124
13	Strack et al. (1988)[36]	17, 1894	0.016	0.041	0.394	0.693	-0.064, 0.096	-0.099, 0.131	Yes	0.115	0.150
14	Srull & Wyer (1979)[34]	26, 7373	0.033	0.014	2.277	0.023	0.006, 0.060	-0.006, 0.072	Yes	0.039	0.051
15	Mazar et al. (2008)[14]	25, 5786	-0.04	0.041	-0.976	0.327	-0.120, 0.040	-0.155, 0.075	Yes	0.115	0.150

Notes: The effect size of the replication study corresponding to the original study by Schwarz et al. (1991)[28] is denoted in Cohen's q units, rather than Cohen's d units. For Hauser et al. (2007)[6] there are two relevant replication estimates as both the replication and the meta-analysis looks at both the difference between scenario 1 and 2 and the difference between scenario 3 and 4. We have pooled these two estimates, since the replication estimates for Hauser et al. (2007)[6] belongs to different slates of Klein et al. (2018)[2], making the two estimates statistically independent. The two replication estimates reported in Alogna et al. (2014)[26] for Schooler & Engstler-Schooler (1990) [25] has been pooled as they are statistically independent and since they are both included in the meta-analysis.

Supplementary Table 5. Effect sizes, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for results reported in Fig. 2, panel b. We also report the smallest detectable effect size difference with 80% power, denoted by "MDE (5%)" and "MDE (0.5%)" for the 15 meta-replication differences (for both the 0.5% and the 5% significance levels).

Original study	Difference	Standard error	z-value	p-value	95% CI	99.5% CI	MDE (5%)	MDE (0.5%)
Oppenheimer et al. (2009) [20]	0.226	0.076	2.965	0.003	0.077, 0.375	0.013, 0.439	0.213	0.277
Tversky & Kahneman (1981) [39]	-0.03	0.041	-0.729	0.466	-0.050, 0.110	-0.085, 0.145	0.115	0.150
Husnu & Crisp (2010) [8]	0.329	0.075	4.365	< 0.001	0.182, 0.476	0.118, 0.540	0.210	0.274
Schwarz et al. (1991) [28]	0.197	0.068	2.903	0.004	0.064, 0.330	0.006, 0.388	0.190	0.248
Hauser et al. (2007) [6]	0.092	0.134	0.684	0.494	-0.171, 0.355	-0.284, 0.468	0.375	0.489
Critcher & Gilovich (2008) [1]	0.27	0.076	3.560	< 0.001	0.121, 0.419	0.057, 0.483	0.213	0.277
Graham, Haidt & Nosek (2009) [4]	0.386	0.053	7.283	< 0.001	0.282, 0.490	0.237, 0.535	0.148	0.193
Jostmann et al. (2009) [11]	0.54	0.067	8.064	< 0.001	0.409, 0.671	0.352, 0.728	0.188	0.245
Monin & Miller (2001) [18]	0.163	0.042	3.869	< 0.001	0.081, 0.245	0.045, 0.281	0.118	0.153
Schooler & Engstler-Schooler (1990) [25]	0.16	0.048	3.352	< 0.001	0.066, 0.254	0.025, 0.295	0.134	0.175
Sripada et al. (2014) [30]	0.58	0.062	9.408	< 0.001	0.458, 0.702	0.406, 0.754	0.174	0.226
Rand et al. (2012) [22]	0.076	0.038	1.988	0.047	0.002, 0.150	-0.031, 0.183	0.106	0.139
Strack et al. (1988) [36]	0.184	0.051	3.621	< 0.001	0.084, 0.284	0.041, 0.327	0.143	0.186
Srull & Wyer (1979) [34]	0.321	0.033	9.718	< 0.001	0.256, 0.386	0.228, 0.414	0.092	0.120
Mazar et al. (2008)[14]	0.47	0.08	5.87	< 0.001	0.313, 0.627	0.245, 0.695	0.224	0.292

Notes: The effect size of the meta-replication difference corresponding to the original study by Schwarz et al. (1991)[28] is denoted in Cohen's q units, rather than Cohen's d units.

Supplementary Table 6. Mean Random Effects difference, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for results reported in Fig. 3. We also report the mean unweighted difference and the smallest detectable mean effect size difference with 80% power, denoted by "MDE (5%)" and "MDE (0.5%)" for the different robustness tests and sub-group analyses (for both the 0.5% and the 5% significance levels).

Test	Mean random effects difference	SE	Z-value	p-value	95% CI	99.5% CI	Mean unweighted difference	MDE 5%	MDE 0.5%
Mean meta-replication effect size difference N=15	0.263	0.045	5.810	<0.001	0.175, 0.352	0.136, 0.391	0.264	0.127	0.165
Excluding non-published meta-analyses N=13	0.253	0.050	5.066	<0.001	0.155, 0.351	0.113, 0.394	0.254	0.140	0.182
Excluding meta-analysis that included the replication N=14	0.270	0.049	5.521	<0.001	0.174, 0.365	0.133, 0.407	0.270	0.137	0.178
Excluding meta-analyses without the original study N=11	0.219	0.049	4.456	<0.001	0.123, 0.315	0.081, 0.357	0.223	0.138	0.179
Including the alternative meta-analysis for Sripada et al. N=15	0.250	0.041	6.145	<0.001	0.170, 0.329	0.136, 0.364	0.252	0.114	0.148
Including alternative meta-analysis for Graham (Loyalty). N=15	0.275	0.049	5.610	<0.001	0.179, 0.372	0.138, 0.413	0.276	0.137	0.179
Including alternative meta-analysis for Graham (Sanctity). N=15	0.262	0.047	5.536	<0.001	0.169, 0.355	0.129, 0.395	0.276	0.133	0.173
Excluding studies using Cohen's q N=14	0.268	0.048	5.534	<0.001	0.173, 0.363	0.132, 0.404	0.269	0.136	0.177
Many Labs replication studies N=9	0.243	0.058	4.211	<0.001	0.130, 0.356	0.081, 0.405	0.241	0.162	0.211
Registered Replication Report replication studies N=6	0.293	0.079	3.717	<0.001	0.139, 0.448	0.072, 0.514	0.299	0.221	0.288
Replications that replicated ($p < 0.05$). N=8	0.209	0.051	4.071	<0.001	0.108, 0.310	0.065, 0.353	0.206	0.144	0.187
Replications that did not replicate ($p < 0.05$). N=7	0.327	0.076	4.314	<0.001	0.178, 0.476	0.114, 0.540	0.331	0.212	0.277
Replications that replicated ($p < 0.005$). N=7	0.190	0.056	3.373	0.001	0.080, 0.301	0.032, 0.348	0.189	0.158	0.206
Replications that did not replicate ($p < 0.005$). N=8	0.326	0.065	5.004	<0.001	0.198, 0.453	0.143, 0.508	0.330	0.182	0.237

Supplementary Table 7. Results from a random-effects meta-regression of the difference measure on a constant term and a dummy for whether the replication successfully detected an effect in the direction of the original study.

	(1)	(2)
Successful replication (p = 0.05)	-0.117 (0.090) [0.213] -0.293, 0.059 {-0.370, 0.136}	
Successful replication (p = 0.005)		-0.134 (0.088) [0.149] -0.306, 0.038 {-0.381, 0.113}
Studies that failed to replicate	0.326 (0.065) [<0.001.] 0.199, 0.453 {0.144, 0.508}	0.325 (0.059) [< 0.001] 0.209, 0.441 {0.159, 0.491}
N	15	15

Notes: Standard errors in parentheses, *p*-values in brackets. 95% Confidence Intervals (CIs) appear below the p-value bracket, and 99.55% CIs appear in curly brackets below the 95% CIs. The first column defines "replication success" as the replication detecting an effect at p = 0.05 in the same direction as the original study and the second column defines "replication success" as the replication detecting an effect at p = 0.005 in the direction of the original study

Supplementary Table 8. Effect sizes, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for results reported in Fig. 4 (the difference in effect size between the original study and the meta-analysis). We also report the smallest detectable effect size difference with 80% power, denoted by "MDE (5%)" and "MDE (0.5%)" for each of the 14 original-meta pairs and the mean original-meta effect size difference (for both the 0.5% and the 5% significance levels).

Original study	Difference	Standard error	z-value (p-value)	95% CI	99.5% CI	MDE (5%)	MDE (0.5%)
Oppenheimer et al. (2009)[20]	-0.266	0.153	-1.735 (0.083)	-0.567, 0.035	-0.696, 0.164	0.429	0.560
Tversky & Kahneman (1981)[39]	0.560	0.124	4.511 (< 0.001)	0.317, 0.803	0.212, 0.908	0.348	0.453
Husnu & Crisp (2010)[8]	0.401	0.371	1.080 (0.280)	-0.327, 1.129	-0.641, 1.443	1.039	1.355
Schwarz et al. (1991)[28]	0.353	0.216	1.632 ()	-0.071, 0.777	-0.254, 0.960	0.606	0.790
Hauser et al. (2007)[6]	0.073	0.149	0.493 (0.622)	-0.219, 0.365	-0.345, 0.492	0.417	0.544
Critcher & Gilovich (2008)[1]	-0.010	0.160	-0.063 (0.950)	-0.323, 0.303	-0.458, 0.438	0.447	0.583
Graham. Haidt & Nosek (2009)[4]	-0.156	0.076	-2.052 (0.040)	-0.304, -0.007	-0.368, 0.057	0.212	0.277
Jostmann et al. (2009)[11]	0.027	0.298	0.091 (0.928)	-0.557, 0.611	-0.809, 0.863	0.834	1.087
Monin & Miller (2001)[18]	0.114	0.155	0.736 (0.462)	-0.190, 0.417	-0.321, 0.549	0.434	0.565
Schooler & Engstler-Schooler (1990)[25]	0.262	0.175	1.497 (0.134)	-0.081, 0.605	-0.229, 0.753	0.490	0.639
Sripada et al. (2014)[30]	-0.330	0.294	-1.121 (0.262)	-0.907, 0.247	-1.157, 0.497	0.825	1.075
Rand et al. (2012)[22]	0.164	0.139	1.179 (0.238)	-0.109, 0.437	-0.227, 0.555	0.390	0.508
Strack et al. (1988)[36]	0.260	0.247	1.053 (0.292)	-0.224, 0.744	-0.433, 0.953	0.691	0.901
Mazar et al. (2008)[14]	0.050	0.208	0.241 (0.810)	-0.357, 0.457	-0.533, 0.633	0.581	0.758
Overall original-meta difference	0.101	0.076	-1.329 (0.184)	-0.048, 0.250	-0.112, 0.314	0.213	0.278

Supplementary Table 9. Effect sizes, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for the random effects meta-analysis results reported in Fig. 5. We also report the smallest detectable effect size with 80% power, denoted by "MDE 0.5%" and "MDE 5%", for each of the 15 meta-replication pairs (for both the 0.5% and the 5% significance levels).

ID	Original study	Estimate	Standard Error	Z-value	p-value	95% CI	99.5% CI	MDE 5%	MDE 0.5%
1	Oppenheimer et al. (2009)[20]	0.501	0.065	7.664	<0.001	0.373, 0.630	0.318, 0.685	0.183	0.239
2	Tversky & Kahneman (1981)[39]	0.532	0.038	13.962	<0.001	0.457, 0.607	0.425, 0.639	0.107	0.139
3	Husnu & Crisp (2010)[8]	0.459	0.069	6.661	<0.001	0.324, 0.594	0.265, 0.652	0.193	0.251
4	Schwarz et al. (1991)[28]	0.128	0.064	2.014	0.044	0.003, 0.253	-0.050, 0.307	0.178	0.232
5	Hauser et al. (2007)[6]	0.887	0.145	6.126	<0.001	0.603, 1.170	0.480, 1.293	0.405	0.528
6	Critcher & Gilovich (2008)[1]	0.304	0.070	4.317	<0.001	0.166, 0.442	0.106, 0.502	0.197	0.257
7	Graham, Haidt & Nosek (2009)[4]	0.670	0.052	12.928	<0.001	0.568, 0.771	0.524, 0.815	0.145	0.189
8	Jostmann et al. (2009)[11]	0.553	0.050	11.112	<0.001	0.455, 0.650	0.413, 0.693	0.139	0.181
9	Monin & Miller (2001)[18]	0.306	0.037	8.237	<0.001	0.233, 0.379	0.202, 0.410	0.104	0.135
10	Schooler & Engstler-Schooler (1990)[25]	0.236	0.066	3.587	<0.001	0.107, 0.365	0.051, 0.421	0.184	0.240
11	Sripada et al. (2014)[30]	0.613	0.024	25.346	<0.001	0.565, 0.660	0.545, 0.680	0.068	0.088
12	Rand et al. (2012)[22]	0.056	0.017	3.384	0.001	0.024, 0.089	0.010, 0.103	0.046	0.060
13	Strack et al. (1988)[36]	0.226	0.026	8.823	<0.001	0.175, 0.276	0.154, 0.297	0.072	0.093
14	Srull & Wyer (1979)[34]	0.382	0.043	8.941	<0.001	0.298, 0.466	0.262, 0.502	0.120	0.156
15	Mazar et al. (2008)[14]	0.426	0.068	6.304	<0.001	0.293, 0.558	0.236, 0.615	0.189	0.246

Supplementary Table 10. Effect sizes, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for the PET-PEESE meta-analysis results reported in Fig. 5. We also report the smallest detectable effect size with 80% power, denoted by "MDE 0.5%" and "MDE 5%", for each of the 15 meta-replication pairs (for both the 0.5% and the 5% significance levels).

ID	Original study	Estimate	Standard Error	Z-value	p-value	95% CI	99.5% CI	MDE 5%	MDE 0.5%
1	Oppenheimer et al. (2009)[20]	0.229	0.195	1.171	0.242	-0.154, 0.612	-0.320, 0.777	0.547	0.713
2	Tversky & Kahneman (1981)[39]	0.548	0.051	10.793	<0.001	0.448, 0.648	0.405, 0.691	0.142	0.185
3	Husnu & Crisp (2010)[8]	-0.386	0.286	-1.347	0.178	-0.947, 0.176	-1.190, 0.418	0.802	1.045
4	Schwarz et al. (1991)[28]	-0.151	0.183	-0.827	0.408	-0.509, 0.207	-0.663, 0.362	0.511	0.666
5	Hauser et al. (2007)[6]	0.513	0.279	1.840	0.066	-0.033, 1.060	-0.270, 1.296	0.781	1.017
6	Critcher & Gilovich (2008)[1]	0.327	0.112	2.910	0.004	0.107, 0.547	0.012, 0.643	0.315	0.410
7	Graham, Haidt & Nosek (2009)[4]	0.713	0.059	12.066	<0.001	0.597, 0.829	0.547, 0.879	0.166	0.216
8	Jostmann et al. (2009)[11]	0.136	0.310	0.440	0.660	-0.471, 0.744	-0.733, 1.006	0.867	1.130
9	Monin & Miller (2001)[18]	-0.053	0.127	-0.416	0.678	-0.303, 0.197	-0.410, 0.305	0.357	0.464
10	Schooler & Engstler-Schooler (1990)[25]	0.293	0.262	1.119	0.263	-0.220, 0.805	-0.441, 1.027	0.732	0.954
11	Sripada et al. (2014)[30]	-0.010	0.061	-0.168	0.866	-0.130, 0.109	-0.181, 0.161	0.171	0.222
12	Rand et al. (2012)[22]	0.008	0.039	0.203	0.839	-0.068, 0.084	-0.101, 0.117	0.109	0.142
13	Strack et al. (1988)[36]	0.015	0.072	0.211	0.833	-0.125, 0.155	-0.186, 0.216	0.200	0.261
14	Srull & Wyer (1979)[34]	0.131	0.136	0.963	0.335	-0.136, 0.399	-0.252, 0.515	0.382	0.498
15	Mazar et al. (2008)[14]	-0.107	0.288	-0.372	0.710	-0.671, 0.457	-0.915, 0.701	0.806	1.049

Supplementary Table 11. Effect sizes, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for the 3PSM meta-analysis results reported in Fig. 5. We also report the smallest detectable effect size with 80% power, denoted by "MDE 0.5%" and "MDE 5%", for each of the 15 meta-replication pairs (for both the 0.5% and the 5% significance levels).

ID	Original study	Estimate	Standard Error	Z-value	p-value	95% CI	99.5% CI	MDE 5%	MDE 0.5%
1	Oppenheimer et al. (2009)[20]	0.298	0.135	2.207	0.027	0.033, 0.562	-0.081, 0.676	0.377	0.492
2	Tversky & Kahneman (1981)[39]	0.543	0.057	9.452	<0.001	0.430, 0.655	0.382, 0.704	0.161	0.209
3	Husnu & Crisp (2010)[8]	0.293	0.103	2.852	0.004	0.092, 0.494	0.005, 0.581	0.288	0.375
4	Schwarz et al. (1991)[28]	0.381	0.243	1.565	0.118	-0.096, 0.858	-0.302, 1.064	0.681	0.887
5	Hauser et al. (2007)[6]	0.846	0.177	4.770	<0.001	0.498, 1.193	0.348, 1.343	0.496	0.647
6	Critcher & Gilovich (2008)[1]	0.231	0.099	2.341	0.019	0.038, 0.424	-0.046, 0.508	0.276	0.360
7	Graham, Haidt & Nosek (2009)[4]	0.663	0.031	21.332	<0.001	0.602, 0.724	0.576, 0.751	0.087	0.113
8	Jostmann et al. (2009)[11]	0.322	0.031	10.347	<0.001	0.261, 0.383	0.235, 0.409	0.087	0.114
9	Monin & Miller (2001)[18]	0.242	0.055	4.379	<0.001	0.134, 0.351	0.087, 0.398	0.155	0.202
10	Schooler & Engstler-Schooler (1990)[25]	0.259	0.062	4.168	<0.001	0.137, 0.381	0.085, 0.434	0.174	0.227
11	Sripada et al. (2014)[30]	0.505	0.039	12.974	<0.001	0.428, 0.581	0.395, 0.614	0.109	0.142
12	Rand et al. (2012)[22]	0.100	0.034	2.972	0.003	0.034, 0.166	0.006, 0.195	0.094	0.123
13	Strack et al. (1988)[36]	0.274	0.048	5.671	<0.001	0.179, 0.369	0.139, 0.410	0.135	0.176
14	Srull & Wyer (1979)[34]	0.426	0.069	6.142	<0.001	0.290, 0.562	0.231, 0.621	0.194	0.253
15	Mazar et al. (2008)[14]	0.418	0.053	7.932	<0.001	0.315, 0.521	0.270, 0.566	0.147	0.192

Supplementary Table 12. Effect sizes, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for the trim and fill meta-analysis results reported in Fig. 5. We also report the smallest detectable effect size with 80% power, denoted by "MDE 0.5%" and "MDE 5%", for each of the 15 meta-replication pairs (for both the 0.5% and the 5% significance levels).

ID	Original study	Estimate	Standard Error	Z-value	p-value	95% CI	99.5% CI	MDE 5%	MDE 0.5%
1	Oppenheimer et al. (2009)[20]	0.501	0.065	7.664	<0.001	0.373, 0.630	0.114, 0.481	0.183	0.239
2	Tversky & Kahneman (1981)[39]	0.576	0.041	13.919	<0.001	0.495, 0.657	0.427, 0.659	0.116	0.151
3	Husnu & Crisp (2010)[8]	0.339	0.074	4.592	<0.001	0.194, 0.483	0.086, 0.500	0.206	0.269
4	Schwarz et al. (1991)[28]	0.128	0.064	2.015	0.044	0.003, 0.253	0.202, 0.559	0.178	0.232
5	Hauser et al. (2007)[6]	0.887	0.145	6.126	<0.001	0.603, 1.170	0.439, 1.252	0.405	0.528
6	Critcher & Gilovich (2008)[1]	0.304	0.070	4.317	<0.001	0.166, 0.442	0.033, 0.429	0.197	0.257
7	Graham, Haidt & Nosek (2009)[4]	0.722	0.047	15.308	<0.001	0.630, 0.815	0.531, 0.796	0.132	0.172
8	Jostmann et al. (2009)[11]	0.513	0.045	11.401	<0.001	0.425, 0.601	0.196, 0.448	0.126	0.164
9	Monin & Miller (2001)[18]	0.224	0.038	5.942	<0.001	0.150, 0.297	0.137, 0.348	0.105	0.137
10	Schooler & Engstler-Schooler (1990)[25]	0.236	0.066	3.587	<0.001	0.107, 0.365	0.075, 0.444	0.184	0.240
11	Sripada et al. (2014)[30]	0.483	0.027	17.922	<0.001	0.430, 0.535	0.429, 0.580	0.075	0.098
12	Rand et al. (2012)[22]	0.056	0.017	3.384	0.001	0.024, 0.089	0.054, 0.147	0.046	0.060
13	Strack et al. (1988)[36]	0.226	0.026	8.823	<0.001	0.175, 0.276	0.203, 0.346	0.072	0.093
14	Srull & Wyer (1979)[34]	0.382	0.043	8.941	<0.001	0.298, 0.466	0.306, 0.546	0.120	0.156
15	Mazar et al. (2008)[14]	0.302	0.074	4.084	<0.001	0.157, 0.447	0.210, 0.626	0.207	0.270

Supplementary Table 13. Effect sizes, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for the random effects meta-analysis results reported in Fig. 6. We also report the smallest detectable effect size with 80% power, denoted by "MDE 0.5%" and "MDE 5%", for each of the 15 meta-replication pairs (for both the 0.5% and the 5% significance levels).

ID	Original study	Difference	Standard Error	Z-value	p-value	95% CI	99.5% CI	MDE 5%	MDE 0.5%
1	Oppenheimer et al. (2009)[20]	0.231	0.075	3.105	0.002	0.085, 0.378	0.022, 0.441	0.209	0.272
2	Tversky & Kahneman (1981)[39]	-0.068	0.052	-1.301	0.193	-0.170, 0.034	-0.215, 0.079	0.146	0.191
3	Husnu & Crisp (2010)[8]	0.329	0.075	4.362	<0.001	0.181, 0.477	0.117, 0.540	0.211	0.275
4	Schwarz et al. (1991)[28]	0.198	0.069	2.891	0.004	0.064, 0.333	0.006, 0.391	0.192	0.250
5	Hauser et al. (2007)[6]	0.102	0.155	0.657	0.511	-0.202, 0.405	-0.333, 0.536	0.433	0.565
6	Critcher & Gilovich (2008)[1]	0.264	0.075	3.525	<0.001	0.117, 0.411	0.054, 0.475	0.210	0.274
7	Graham, Haidt & Nosek (2009)[4]	0.380	0.057	6.700	<0.001	0.269, 0.491	0.221, 0.539	0.159	0.207
8	Jostmann et al. (2009)[11]	0.523	0.066	7.922	<0.001	0.394, 0.652	0.338, 0.708	0.185	0.241
9	Monin & Miller (2001)[18]	0.159	0.041	3.866	<0.001	0.079, 0.240	0.044, 0.275	0.115	0.150
10	Schooler & Engstler-Schooler (1990)[25]	0.160	0.067	2.380	0.017	0.028, 0.292	-0.029, 0.349	0.188	0.245
11	Sripada et al. (2014)[30]	0.573	0.061	9.371	<0.001	0.453, 0.692	0.401, 0.744	0.171	0.223
12	Rand et al. (2012)[22]	0.076	0.038	1.992	0.046	0.001, 0.151	-0.031, 0.183	0.107	0.140
13	Strack et al. (1988)[36]	0.210	0.048	4.371	<0.001	0.116, 0.304	0.075, 0.344	0.134	0.175
14	Srull & Wyer (1979)[34]	0.349	0.045	7.737	<0.001	0.261, 0.438	0.223, 0.476	0.126	0.165
15	Mazar et al. (2008)[14]	0.466	0.079	5.901	<0.001	0.311, 0.620	0.244, 0.687	0.221	0.288

Supplementary Table 14. Effect sizes, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for the PET-PEESE meta-analysis results reported in Fig. 6. We also report the smallest detectable effect size with 80% power, denoted by "MDE 0.5%" and "MDE 5%", for each of the 15 meta-replication pairs (for both the 0.5% and the 5% significance levels).

ID	Original study	Difference	Standard Error	Z-value	p-value	95% CI	99.5% CI	MDE 5%	MDE 0.5%
1	Oppenheimer et al. (2009)[20]	-0.041	0.199	-0.208	0.835	-0.431, 0.348	-0.599, 0.516	0.556	0.725
2	Tversky & Kahneman (1981)[39]	-0.052	0.062	-0.838	0.402	-0.174, 0.070	-0.226, 0.122	0.174	0.227
3	Husnu & Crisp (2010)[8]	-0.516	0.288	-1.790	0.073	-1.080, 0.049	-1.324, 0.293	0.807	1.052
4	Schwarz et al. (1991)[28]	-0.081	0.184	-0.439	0.661	-0.442, 0.280	-0.598, 0.437	0.516	0.673
5	Hauser et al. (2007)[6]	-0.272	0.284	-0.956	0.339	-0.829, 0.285	-1.070, 0.526	0.796	1.037
6	Critcher & Gilovich (2008)[1]	0.287	0.115	2.491	0.013	0.061, 0.513	-0.036, 0.611	0.323	0.421
7	Graham, Haidt & Nosek (2009)[4]	0.423	0.063	6.674	<0.001	0.299, 0.547	0.245, 0.601	0.178	0.231
8	Jostmann et al. (2009)[11]	0.106	0.313	0.340	0.734	-0.507, 0.719	-0.772, 0.984	0.876	1.142
9	Monin & Miller (2001)[18]	-0.200	0.129	-1.552	0.121	-0.452, 0.053	-0.560, 0.161	0.360	0.469
10	Schooler & Engstler-Schooler (1990)[25]	0.217	0.262	0.827	0.408	-0.297, 0.730	-0.518, 0.952	0.733	0.956
11	Sripada et al. (2014)[30]	-0.050	0.083	-0.607	0.544	-0.213, 0.112	-0.283, 0.182	0.232	0.302
12	Rand et al. (2012)[22]	0.028	0.052	0.537	0.591	-0.074, 0.130	-0.118, 0.174	0.146	0.190
13	Strack et al. (1988)[36]	-0.001	0.082	-0.011	0.991	-0.162, 0.160	-0.232, 0.230	0.230	0.300
14	Srull & Wyer (1979)[34]	0.098	0.137	0.717	0.473	-0.171, 0.368	-0.287, 0.484	0.384	0.501
15	Mazar et al. (2008)[14]	-0.067	0.291	-0.231	0.818	-0.637, 0.502	-0.883, 0.749	0.814	1.061

Supplementary Table 15. Effect sizes, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for the 3PSM meta-analysis results reported in Fig. 6. We also report the smallest detectable effect size with 80% power, denoted by "MDE 0.5%" and "MDE 5%", for each of the 15 meta-replication pairs (for both the 0.5% and the 5% significance levels).

ID	Original study	Difference	Standard Error	Z-value	p-value	95% CI	99.5% CI	MDE 5%	MDE 0.5%
1	Oppenheimer et al. (2009)[20]	0.028	0.139	0.197	0.844	-0.246, 0.301	-0.364, 0.419	0.390	0.509
2	Tversky & Kahneman (1981)[39]	-0.057	0.068	-0.847	0.397	-0.190, 0.075	-0.247, 0.133	0.189	0.247
3	Husnu & Crisp (2010)[8]	0.163	0.107	1.520	0.129	-0.047, 0.373	-0.138, 0.464	0.300	0.391
4	Schwarz et al. (1991)[28]	0.451	0.245	1.843	0.065	-0.029, 0.930	-0.236, 1.137	0.685	0.893
5	Hauser et al. (2007)[6]	0.061	0.186	0.327	0.744	-0.303, 0.424	-0.460, 0.581	0.519	0.677
6	Critcher & Gilovich (2008)[1]	0.191	0.102	1.874	0.061	-0.009, 0.391	-0.095, 0.477	0.285	0.372
7	Graham, Haidt & Nosek (2009)[4]	0.373	0.039	9.660	<0.001	0.298, 0.449	0.265, 0.482	0.108	0.141
8	Jostmann et al. (2009)[11]	0.292	0.053	5.472	<0.001	0.187, 0.397	0.142, 0.442	0.149	0.195
9	Monin & Miller (2001)[18]	0.096	0.058	1.646	0.100	-0.018, 0.210	-0.068, 0.259	0.163	0.212
10	Schooler & Engstler-Schooler (1990)[25]	0.183	0.064	2.877	0.004	0.058, 0.308	0.004, 0.362	0.179	0.233
11	Sripada et al. (2014)[30]	0.465	0.068	6.804	<0.001	0.331, 0.598	0.273, 0.656	0.191	0.249
12	Rand et al. (2012)[22]	0.120	0.048	2.495	0.013	0.026, 0.215	-0.015, 0.256	0.135	0.176
13	Strack et al. (1988)[36]	0.258	0.063	4.092	<0.001	0.135, 0.382	0.081, 0.435	0.177	0.230
14	Srull & Wyer (1979)[34]	0.393	0.071	5.547	<0.001	0.254, 0.532	0.194, 0.592	0.198	0.259
15	Mazar et al. (2008)[14]	0.458	0.067	6.870	<0.001	0.327, 0.588	0.271, 0.645	0.187	0.243

Supplementary Table 16. Effect sizes, standard errors, z-value, p-value and confidence interval (95% and 99.5%) for the trim and fill meta-analysis results reported in Fig. 6. We also report the smallest detectable effect size with 80% power, denoted by "MDE 0.5%" and "MDE 5%", for each of the 15 meta-replication pairs (for both the 0.5% and the 5% significance levels).

ID	Original study	Difference	Standard Error	Z-value	p-value	95% CI	99.5% CI	MDE 5%	MDE 0.5%
1	Oppenheimer et al. (2009)[20]	0.231	0.075	3.105	0.002	0.085, 0.378	0.022, 0.441	0.209	0.272
2	Tversky & Kahneman (1981)[39]	-0.024	0.055	-0.447	0.655	-0.132, 0.083	-0.178, 0.129	0.153	0.199
3	Husnu & Crisp (2010)[8]	0.209	0.080	2.613	0.009	0.052, 0.365	-0.015, 0.433	0.224	0.291
4	Schwarz et al. (1991)[28]	0.198	0.069	2.891	0.004	0.064, 0.333	0.006, 0.391	0.192	0.250
5	Hauser et al. (2007)[6]	0.102	0.155	0.657	0.511	-0.202, 0.405	-0.333, 0.536	0.433	0.565
6	Critcher & Gilovich (2008)[1]	0.264	0.075	3.525	<0.001	0.117, 0.411	0.054, 0.475	0.210	0.274
7	Graham, Haidt & Nosek (2009)[4]	0.432	0.052	8.237	<0.001	0.329, 0.535	0.285, 0.579	0.147	0.191
8	Jostmann et al. (2009)[11]	0.483	0.062	7.729	<0.001	0.361, 0.605	0.308, 0.658	0.175	0.228
9	Monin & Miller (2001)[18]	0.077	0.042	1.851	0.064	-0.005, 0.159	-0.040, 0.194	0.117	0.152
10	Schooler & Engstler-Schooler (1990)[25]	0.160	0.067	2.380	0.017	0.028, 0.292	-0.029, 0.349	0.188	0.245
11	Sripada et al. (2014)[30]	0.443	0.062	7.109	<0.001	0.321, 0.565	0.268, 0.617	0.174	0.227
12	Rand et al. (2012)[22]	0.076	0.038	1.992	0.046	0.001, 0.151	-0.031, 0.183	0.107	0.140
13	Strack et al. (1988)[36]	0.210	0.048	4.371	<0.001	0.116, 0.304	0.075, 0.344	0.134	0.175
14	Srull & Wyer (1979)[34]	0.349	0.045	7.737	<0.001	0.261, 0.438	0.223, 0.476	0.126	0.165
15	Mazar et al. (2008)[14]	0.342	0.085	4.050	<0.001	0.177, 0.508	0.105, 0.580	0.237	0.308

Supplementary references

1 Critcher, C. R. & Gilovich, T. Incidental environmental anchors. *Journal of Behavioral Decision Making* **21**, 241-251 (2008).

2 Klein, R. A. *et al.* Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods Practices in Psychological Science* **1**, 443-490 (2018).

3 Henriksson, K. A. C. *Irrelevant quantity effects: a meta-analysis*, (2015).

4 Graham, J., Haidt, J. & Nosek, B. A. Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality Social Psychology* **96**, 1029 (2009).

5 Kivikangas, J. M., Lönnqvist, J.-E. & Ravaja, N. Relationships between Moral Foundations and political orientation–local study and meta-analysis. Poster presented at Annual Convention of Society for Personality and Social Psychology, San Diego, January 2016. DOI: 10.13140/RG.2.1.2277.0964.

6 Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R. & Mikhail, J. A dissociation between moral judgments and justifications. *Mind Language* **22**, 1-21 (2007).

7 Feltz, A. & May, J. The means/side-effect distinction in moral cognition: A meta-analysis. *Cognition* **166**, 314-327 (2017).

8 Husnu, S. & Crisp, R. J. Elaboration enhances the imagined contact effect. *Journal of Experimental Social Psychology* **46**, 943-950 (2010).

9 Klein, R. A. *et al.* Investigating variation in replicability. *Social Psychology* (2014).

10 Miles, E. & Crisp, R. J. A meta-analytic test of the imagined contact hypothesis. *Group Processes & Intergroup Relations* **17**, 3-26 (2014).

11 Jostmann, N. B., Lakens, D. & Schubert, T. W. Weight as an embodiment of importance. *Psychological Science* **20**, 1169-1174 (2009).

12 Ebersole, C. R. *et al.* Many Labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology* **67**, 68-82 (2016).

13 Rabelo, A. L., Keller, V. N., Pilati, R. & Wicherts, J. M. No effect of weight on judgments of importance in the moral domain and evidence of publication bias from a meta-analysis. *PloS one* **10**, e0134808 (2015).

14 Mazar, N., Amir, O. & Ariely, D. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research* **45**, 633-644 (2008).

15 Verschuere, B. *et al.* Registered replication report on Mazar, Amir, and Ariely (2008). *Advances in Methods and Practices in Psychological Science* **1**, 299-317 (2018).

16 McCarthy, R. J. *et al.* Registered Replication Report on Srull and Wyer (1979). *Advances in Methods Practices in Psychological Science* **1**, 321-336 (2018).

17 Belle, N. & Cantarelli, P. What Causes Unethical Behavior? A Meta-Analysis to Set an Agenda for Public Administration Research. *Public Administration Review* **77**, 327-339 (2017).

18 Monin, B. & Miller, D. T. Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology* **81**, 33 (2001).

19 Blanken, I., van de Ven, N. & Zeelenberg, M. A meta-analytic review of moral licensing. *Personality and Social Psychology Bulletin* **41**, 540-558 (2015).

20 Oppenheimer, D. M., Meyvis, T. & Davidenko, N. Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology* **45**, 867-872 (2009).

21 Roth, S., Robbert, T. & Straus, L. On the sunk-cost effect in economic decision-making: a meta-analytic review. *Business Research* **8**, 99-138 (2015).

22 Rand, D. G., Greene, J. D. & Nowak, M. A. Spontaneous giving and calculated greed. *Nature* **489**, 427 (2012).

23 Bouwmeester, S. *et al.* Registered Replication Report: Rand, Greene, and Nowak (2012). *Perspectives on Psychological Science* **12**, 527-542 (2017).

24 Rand, D. G. Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science* **27**, 1192-1206 (2016).

25 Schooler, J. W. & Engstler-Schooler, T. Y. Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology* **22**, 36-71 (1990).

26 Alogna, V. *et al.* Registered Replication Report: Schooler and Engstler-Schooler (1990). *Perspectives on Psychological Science* **9**, 556-578 (2014).

27 Meissner, C. A. & Brigham, J. C. A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition* **15**, 603-616 (2001).

28 Schwarz, N., Strack, F. & Mai, H.-P. Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Quarterly* **55**, 3-23 (1991).

29 Schimmack, U. & Oishi, S. The influence of chronically and temporarily accessible information on life satisfaction judgments. *Journal of Personality and Social Psychology* **89**, 395 (2005).

30 Sripada, C., Kessler, D. & Jonides, J. Methylphenidate blocks effort-induced depletion of regulatory control in healthy volunteers. *Psychological Science* **25**, 1227-1234 (2014).

31 Hagger, M. S. *et al.* A Multilab Preregistered Replication of the Ego-Depletion Effect. *Perspectives on Psychological Science* **11**, 546-573 (2016).

32 Hagger, M. S., Wood, C., Stiff, C. & Chatzisarantis, N. L. Ego depletion and the strength model of self-control: a meta-analysis. *Psychological Bulletin* **136**, 495 (2010).

33 Carter, E. C., Kofler, L. M., Forster, D. E. & McCullough, M. E. A series of meta-analytic tests of the depletion effect: self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General* **144**, 796 (2015).

34 Srull, T. K., Wyer, R. S. & psychology, S. The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality* **37**, 1660 (1979).

35 DeCoster, J. & Claypool, H. M. A meta-analysis of priming effects on impression formation supporting a general model of informational biases. *Personality and Social Psychology Review* **8**, 2-27 (2004).

36 Strack, F., Martin, L. L. & Stepper, S. Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology* **54**, 768 (1988).

37 Wagenmakers, E.-J. *et al.* Registered Replication Report: Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science* **11**, 917-928 (2016).

38 Coles, N. A., Larsen, J. T. & Lench, H. C. A meta-analysis of the facial feedback literature: Effects of facial feedback on emotional experience are small and variable. *Psychological Bulletin* (2019).

39 Tversky, A. & Kahneman, D. The framing of decisions and the psychology of choice. *Science* **211**, 453-458 (1981).

40 Kühberger, A. The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior and Human Decision Processes* **75**, 23-55 (1998).

41 Anderson, C., Kraus, M. W., Galinsky, A. D. & Keltner, D. The local-ladder effect: Social status and subjective well-being. *Psychological Science* **23**, 764-771 (2012).

42 Pinquart, M. & Sörensen, S. Influences of socioeconomic status, social network, and competence on subjective well-being in later life: a meta-analysis. *Psychology and Aging* **15**, 187 (2000).

43 Boroditsky, L. Metaphoric structuring: Understanding time through spatial metaphors. *Cognition* **75**, 1-28 (2000).

44 Stickles, E. & Lewis, T. N. Wednesday's Meeting Really Is on Friday: A Meta-Analysis and Evaluation of Ambiguous Spatiotemporal Language. *Cognitive Science* **42**, 1015-1025 (2018).

45 Cacioppo, J. T., Petty, R. E. & Morris, K. J. Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology* **45**, 805 (1983).

46 Von Stumm, S. & Ackerman, P. L. Investment and intellect: A review and meta-analysis. *Psychological Bulletin* **139**, 841 (2013).

47 Stiff, J. B. Cognitive processing of persuasive message cues: A meta-analytic review of the effects of supporting information on attitudes. *Communications Monographs* **53**, 75-89 (1986).

48 Caruso, E. M., Vohs, K. D., Baxter, B. & Waytz, A. Mere exposure to money increases endorsement of free-market systems and social inequality. *Journal of Experimental Psychology: General* **142**, 301 (2013).

49 Vadillo, M. A., Hardwicke, T. E. & Shanks, D. R. Selection bias, vote counting, and money-priming effects: A comment on Rohrer, Pashler, and Harris (2015) and Vohs (2015). *Journal of Experimental Psychology: General* **145**, 655-663 (2016).

50 Schuler, J. & Wänke, M. A fresh look on money priming: Feeling privileged or not makes a difference. *Social Psychological and Personality Science* **7**, 366-373 (2016).

51 Finkel, E. J., Rusbult, C. E., Kumashiro, M. & Hannon, P. A. Dealing with betrayal in close relationships: Does commitment promote forgiveness? *Journal of Personality and Social Psychology* **82**, 956 (2002).

52 Cheung, I. *et al.* Registered Replication Report: Study 1 from Finkel, Rusbult, Kumashiro, & Hannon (2002). *Perspectives on Psychological Science* **11**, 750-764 (2016).

53 Fehr, R., Gelfand, M. J. & Nag, M. The road to forgiveness: a meta-analytic synthesis of its situational and dispositional correlates. *Psychological Bulletin* **136**, 894 (2010).

54 Galinsky, A. D., Magee, J. C., Inesi, M. E. & Gruenfeld, D. H. Power and perspectives not taken. *Psychological Science* **17**, 1068-1074 (2006).

55 Hall, J. A., Mast, M. S. & Latu, I.-M. The vertical dimension of social relations and accurate interpersonal perception: A meta-analysis. *Journal of Nonverbal Behavior* **39**, 131-163 (2015).

56 Skitka, L. J., Morgan, G. S. & Wisneski, D. C. in *Social psychology and politics* (eds Joseph P Forgas, Klaus Fiedler, & William D Crano) 73-90 (Psychology Press, 2015).

57 Inbar, Y., Pizarro, D. A., Knobe, J. & Bloom, P. Disgust sensitivity predicts intuitive disapproval of gays. *Emotion* **9**, 435 (2009).

58 Terrizzi Jr, J. A., Shook, N. J. & McDaniel, M. A. The behavioral immune system and social conservatism: A meta-analysis. *Evolution and Human Behavior* **34**, 99-108 (2013).

59 Kiss, M. J., Morrison, M. A. & Morrison, T. G. A Meta-Analytic Review of the Association Between Disgust and Prejudice Toward Gay Men. *Journal of Homosexuality*, 1-23 (2018).

60 Jacowitz, K. E. & Kahneman, D. Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin* **21**, 1161-1166 (1995).

61 Orr, D. & Guthrie, C. Anchoring, information, expertise, and negotiation: New insights from meta-analysis. *Ohio St. J. on Disp. Resol.* **21**, 597 (2005).

62 Nosek, B. A., Banaji, M. R. & Greenwald, A. G. Math= male, me= female, therefore math≠ me. *Journal of Personality and Social Psychology* **83**, 44 (2002).

63 Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L. & Banaji, M. R. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology* **97**, 17 (2009).

64 Hofmann, W., Gawronski, B., Gschwendner, T., Le, H. & Schmitt, M. A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin* **31**, 1369-1385 (2005).

65 Lindberg, S. M., Hyde, J. S., Petersen, J. L. & Linn, M. C. New trends in gender and mathematics performance: a meta-analysis. *Psychological Bulletin* **136**, 1123 (2010).

66 Reilly, D., Neumann, D. L. & Andrews, G. Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of Educational Progress assessments. *Journal of Educational Psychology* **107**, 645 (2015).

67 Ross, L., Greene, D. & House, P. The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology* **13**, 279-301 (1977).

68 Mullen, B. Strength and immediacy of sources: A meta-analytic evaluation of the forgotten elements of social impact theory. *Journal of Personality and Social Psychology* **48**, 1458 (1985).

69 Rugg, D. Experiments in wording questions: II. *Public Opinion Quarterly* **5**, 91 (1941).

70 Holleman, B. Wording effects in survey research using meta-analysis to explain the forbid/allow asymmetry. *Journal of Quantitative Linguistics* **6**, 29-40 (1999).

71 Meissner, C. A., Sporer, S. L. & Susa, K. J. A theoretical review and meta-analysis of the description-identification relationship in memory for faces. *European Journal of Cognitive Psychology* **20**, 414-455 (2008).

72 Heller, D., Watson, D. & Ilies, R. The role of person versus situation in life satisfaction: A critical examination. *Psychological Bulletin* **130**, 574 (2004).

73 Zhong, C.-B. & Liljenquist, K. Washing away your sins: Threatened morality and physical cleansing. *Science* **313**, 1451-1452 (2006).
74 Matsumoto, D. The role of facial response in the experience of emotion: More methodological problems and a meta-analysis. *Journal of Personality and Social Psychology* **52**, 769 (1987).