Appendix A

Table A. *Study characteristics of the included meta-analyses.*

Study	k	N	Intervention	Outcome	Mean effect size (in *d* or *g*)	Uniqueness (in %)
Ahir & Chakraborty, 2021	40	195,103	Feedback	Conservation	*d* = 0.10 [0.09, 0.11]	77%
Bergquist et al., 2019	91	227,730	Social comparison	PEB	*d* = 0.32 [0.28, 0.37]	66%
Karlin et al., 2015	42	256,536	Feedback	Conservation	*d* = 0.14 [0.09, 0.19]	46%
Khanna et al., 2021	360	1,132,864	Financial incentives, education, feedback, social comparison & commitment	Conservation	*d* = 0.30 [0.24, 0.34]	58%
Maki et al., 2016	25	2971	Financial incentives	PEB	*d* = 0.36 [0.22, 0.50]	32%
Mi et al., 2021	112	13998	Financial incentives & education	Energy conservation	*d* = 0.43 [0.37, 0.48]	48%
Nisa et al., 2019	144	3,092,678	Education, social comparison, commitment, appeals & nudges	PEB	*d* = 0.19 [0.15, 0.24]	37%
Osbaldiston & Schott, 2012	253		Various	PEB	*g* = 0.45 [0.43, 0.47]	48%
Semenescu et al., 2020	41	11206	Education & feedback	Transportation behavior	*g* = 0.16 [0.11, 0.21]	58%
Varotto & Spagnolli, 2017	70		Education, feedback, commitment, financial incentives, nudging & social influence	Recycling	*g* = 0.29 [0.24, 0.33]	64%

Appendix B

Formulas for effect size conversion

Converting from *r* to *d* (Borenstein, 2009)

$$d = \frac{2\,r}{\sqrt{1-r^2}} \quad V_d = \frac{4\,V_r}{(1-r^2)^3}$$

Converting from Fisher's *z* scale to *r* (Borenstein, 2009)

$$r = \frac{e^{2z}-1}{e^{2z}+1}$$

Calculating Variance from Confidence Intervals (Schmidt & Hunter, 2015)

$$SE = \frac{(95\%CI_U - 95\%CI_L)}{2 \times 1.96}$$

$$Variance(d) = SE^2$$
$$Variance(r) = SE^2$$

Corrected Covered Area

$$CCA = \frac{N-r}{(r \times c)-r},$$

where *N* is the sum of the number of all primary studies per meta-analysis, r is the number of unique primary studies across meta-analyses and c is the number of meta-analyses (Pieper et al., 2014).

Appendix C

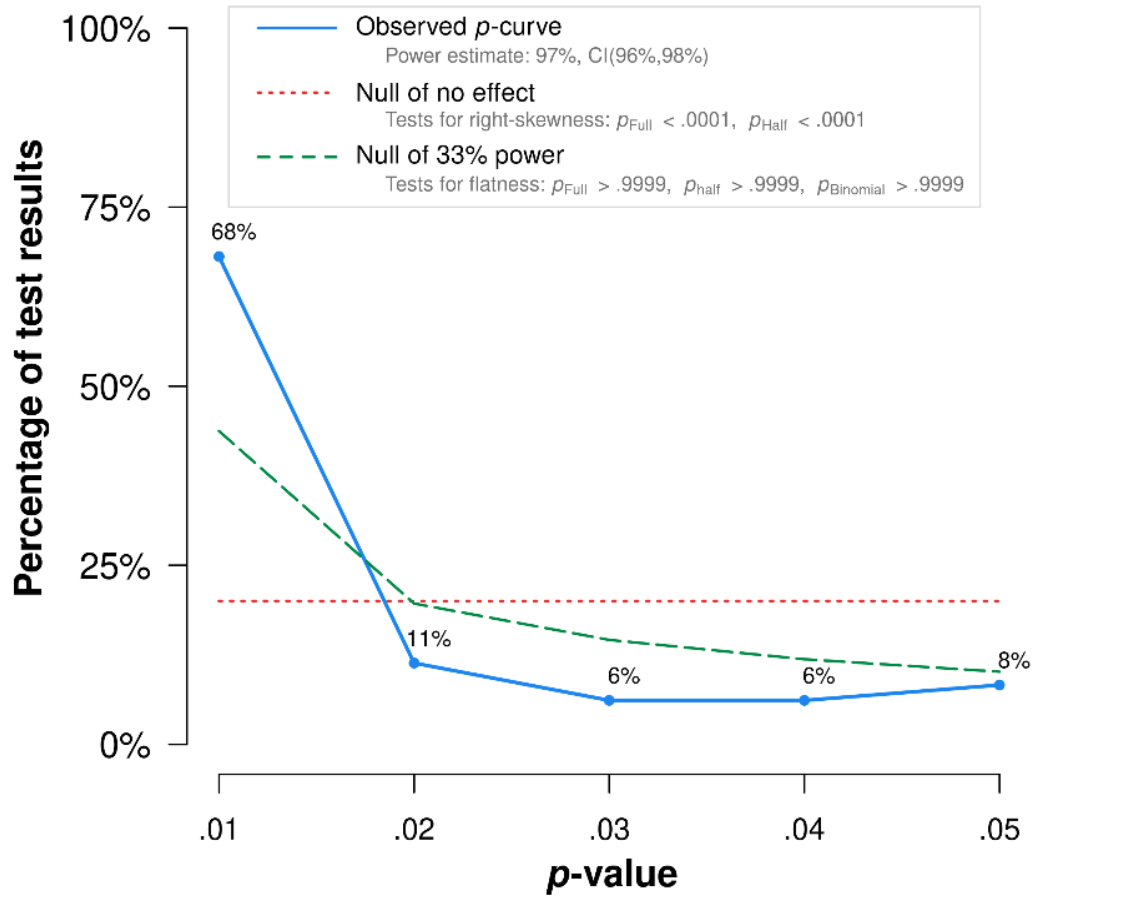
Table B. *Assessment of publication bias in each included meta-analysis.*

Authors and Year	K estimates	Type of assessment	Interpretation (by original authors)
Nisa et al. (2019)	144	Funnel plot, Egger's test	Small-study bias, asymmetry
Mi et al. (2021)	112	Funnel plot, fail-safe N	Low probability of publication bias
Osbaldiston & Schott (2012)	253	None	None
Maki et al. (2016)	25	Egger's test	No small-study bias present
Varotto & Spagnolli (2017)	70	Egger's test, rank correlation test, trim and fill, fail-safe N	Small-study bias, would lower the effect size ($\Delta g = -0.2$), moderate publication bias
Karlin et al. (2015)	42	Moderator analysis on publication type and sample size, trim and fill	Small-study bias, moderate publication bias
Bergquist et al. (2019)	91	Funnel plot, trim and fill, fail-safe N	asymmetry, slightly lower effect size
Semenescu et al. (2020)	41	Funnel plot, trim and fill	Low asymmetry, slightly lower effect size ($\Delta g = -0.053$), moderate publication bias
Khanna et al. (2021)	360	Funnel plot, Egger's test	Small-study bias, Asymmetry
Ahir & Chakraborty (2021)	40	None	None

Note. Asymmetry refers to the fact that the authors detected an asymmetric distribution in the funnel plot.

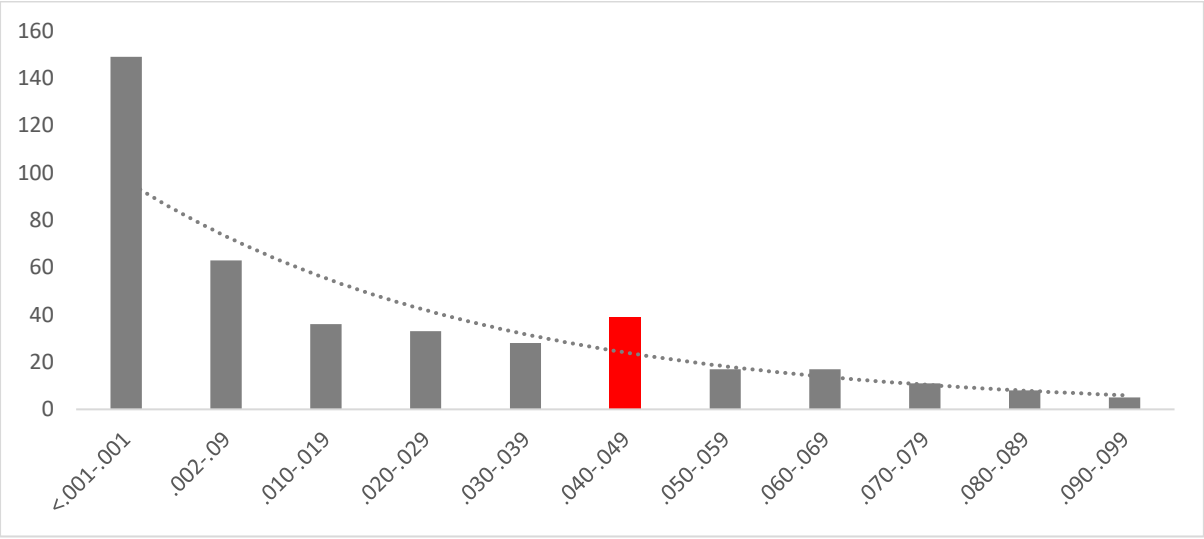
Appendix D

Figure A. *P-curve for 326 studies.*



Note: The observed *p*-curve includes 326 statistically significant ($p < .05$) results, of which 271 are $p < .025$. There were 305 additional results entered but excluded from *p*-curve because they were $p > .05$.

Figure B. *Distributional p-curve for 326 studies.*



Appendix E

Figure D. *Distribution of number of participants included in each type of interventions.*

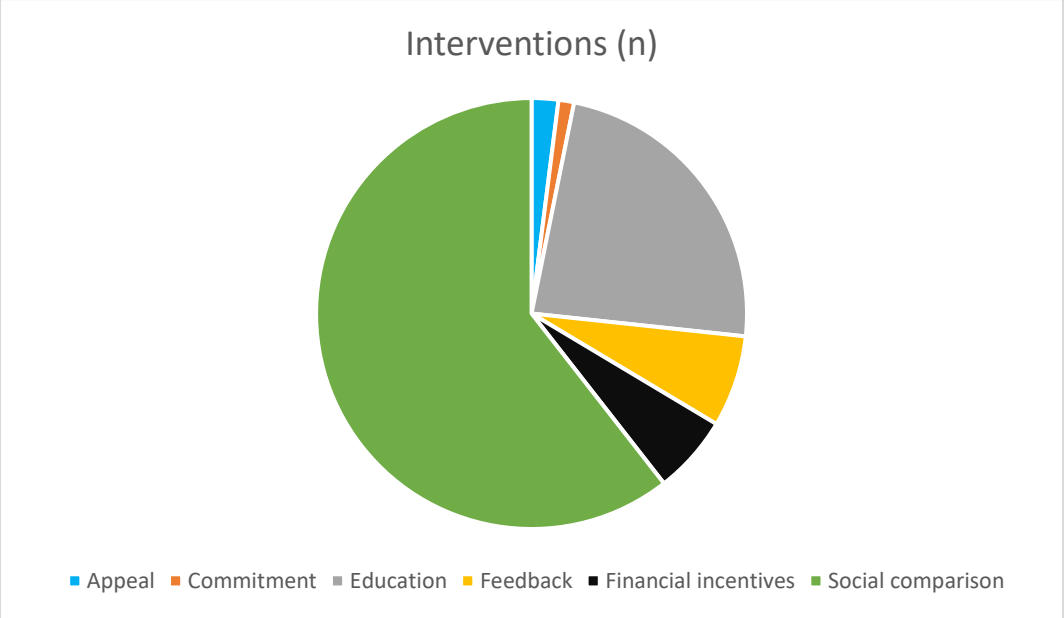


Figure E. *Distribution of type of interventions.*

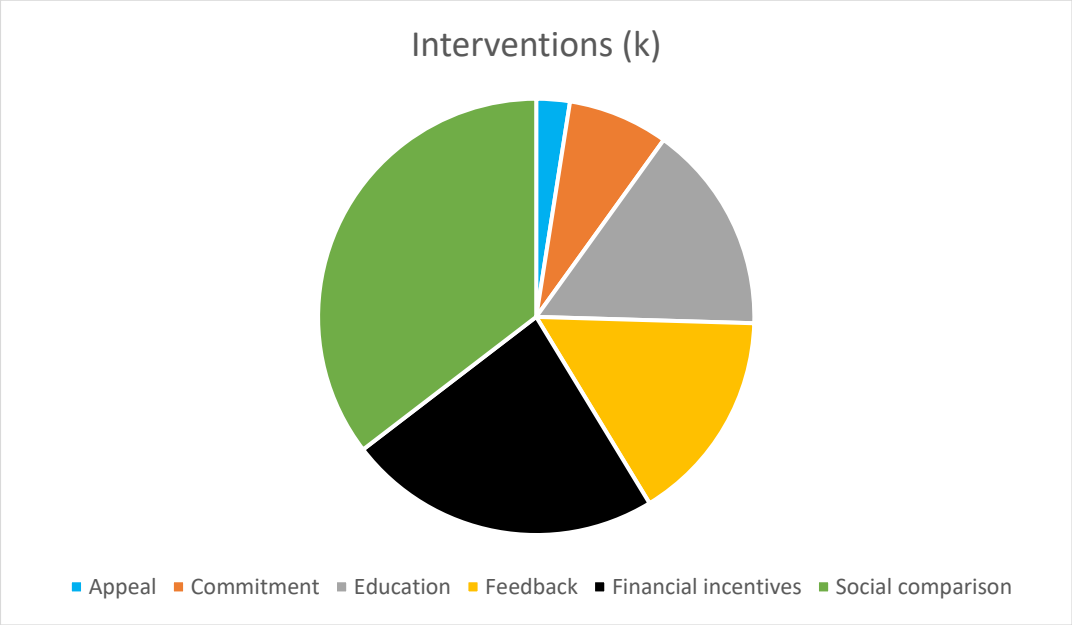
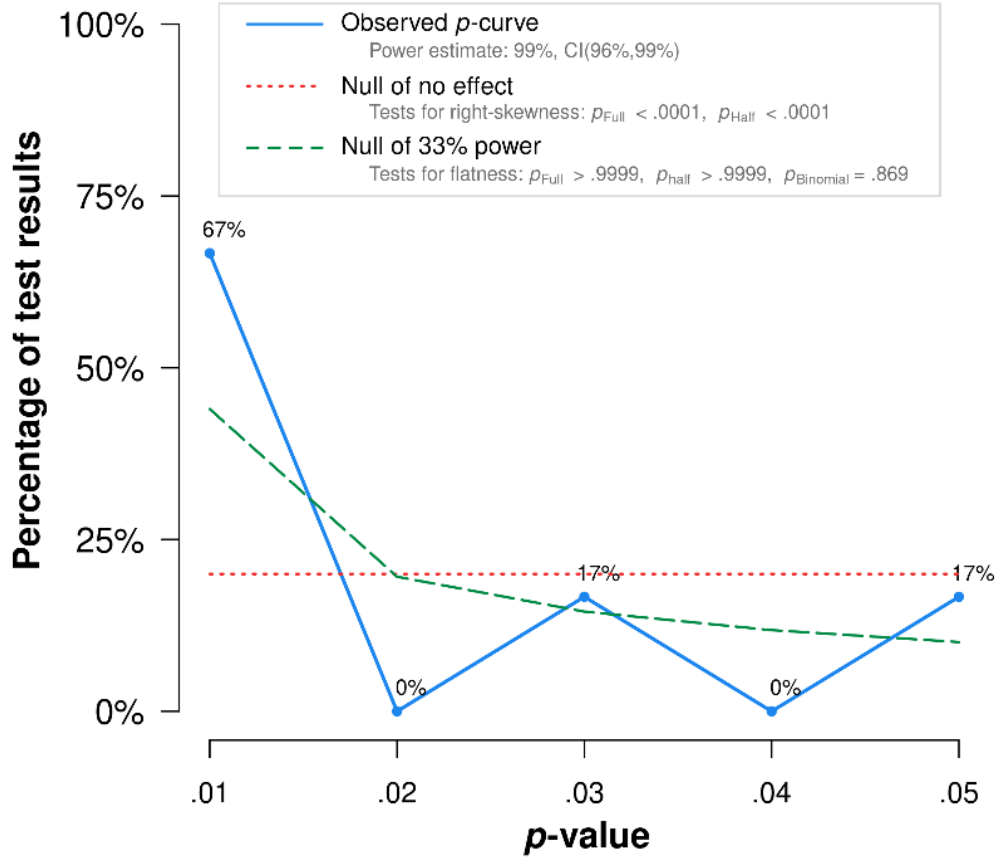


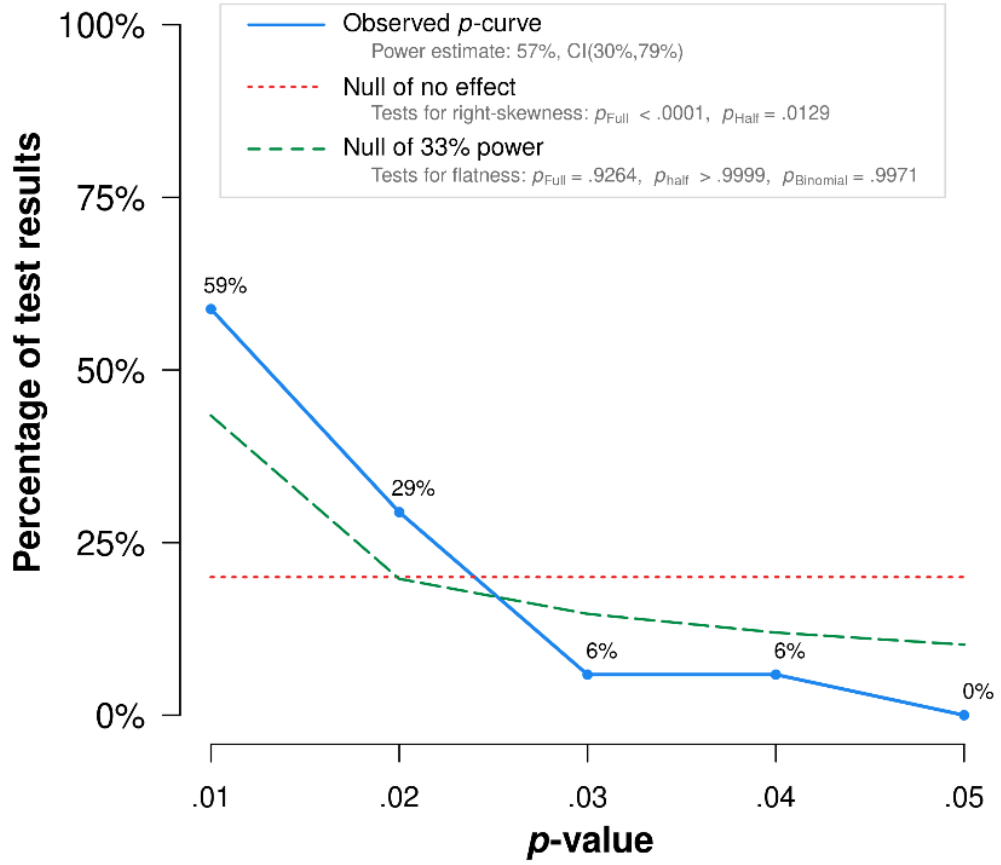
Figure F1. *Descriptive statistics and p-curve from intervention type Appeal*
k = 8
n: min = 100, max = 13969, mean = 2494, sd = 4885, sum = 19956
WAAP: r = .24 (95% CI .13, .34), k = 6
Selection models: N/A
p-curve analysis:



Note: The observed *p*-curve includes 6 statistically significant ($p < .05$) results, of which 5 are $p < .025$.
There were 2 additional results entered but excluded from *p*-curve because they were $p > .05$.

	Binomial Test *(Share of results p<.025)*	**Continuous Test** *(Aggregate with Stouffer Method)*	
		Full p-curve (p's<.05)	**Half p-curve (p's<.025)**
1) Studies contain evidential value. *(Right skew)*	*p*=.1094	*Z*=-6.95, *p*<.0001	*Z*=-7.61, *p*<.0001
2) Studies' evidential value, if any, is inadequate. *(Flatter than 33% power)*	*p*=.869	*Z*=4.66, *p*>.9999	*Z*=6.97, *p*>.9999
	Statistical Power		
Power of tests included in *p*-curve *(correcting for selective reporting)*	Estimate: 99% 90% Confidence interval: (96% , 99%)		

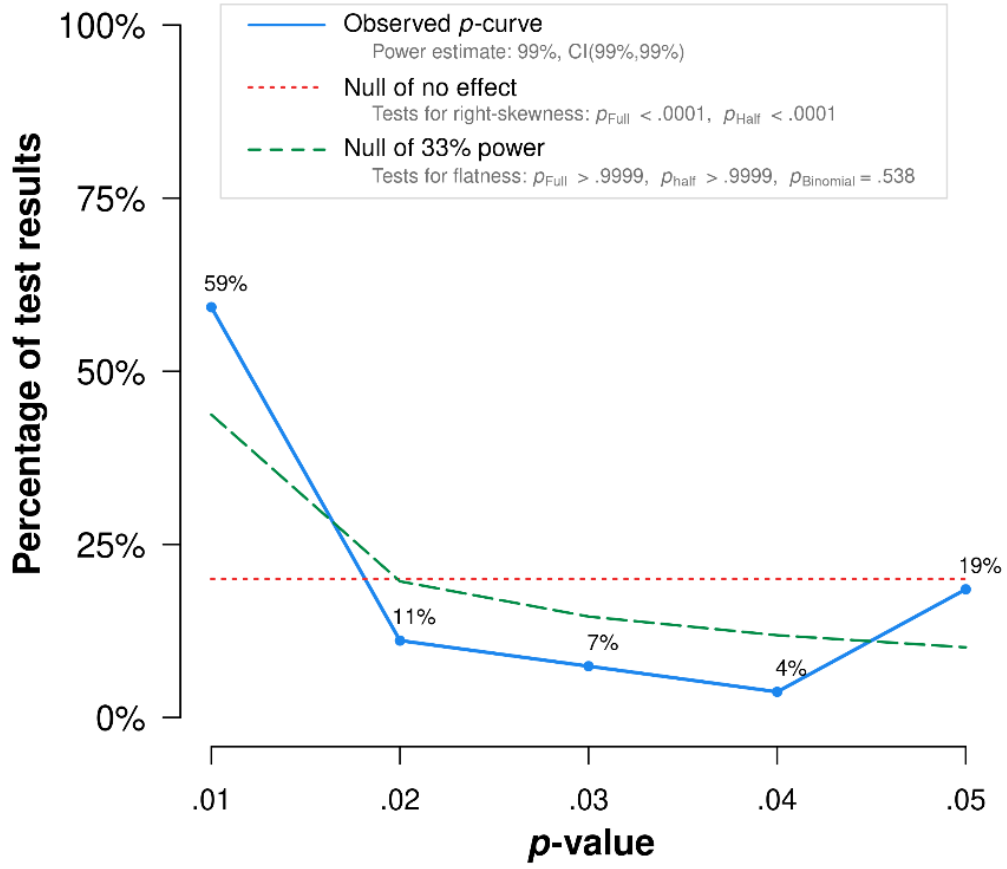
Figure F2. *Descriptive statistics and p-curve from intervention type Commitment*
k = 24
n: min = 11, max = 1791, mean = 282, sd = 694, sum = 11620
WAAP: r =.065 (95% CI.06, 07), k = 5
Selection models: r =.07 (95% CI .05, 09)
p-curve analysis:



Note: The observed *p*-curve includes 17 statistically significant ($p < .05$) results, of which 16 are $p < .025$.
There were 7 additional results entered but excluded from *p*-curve because they were $p > .05$.

	Binomial Test *(Share of results p<.025)*	**Continuous Test** *(Aggregate with Stouffer Method)*	
		Full p-curve (p's<.05)	**Half p-curve (p's<.025)**
1) Studies contain evidential value. *(Right skew)*	*p*=.0001	*Z*=-4.31, *p*<.0001	*Z*=-2.23, *p*=.0129
2) Studies' evidential value, if any, is inadequate. *(Flatter than 33% power)*	*p*=.9971	*Z*=1.45, *p*=.9264	*Z*=4.18, *p*>.9999
	Statistical Power		
Power of tests included in *p*-curve *(correcting for selective reporting)*	Estimate: 57% 90% Confidence interval: (30% , 79%)		

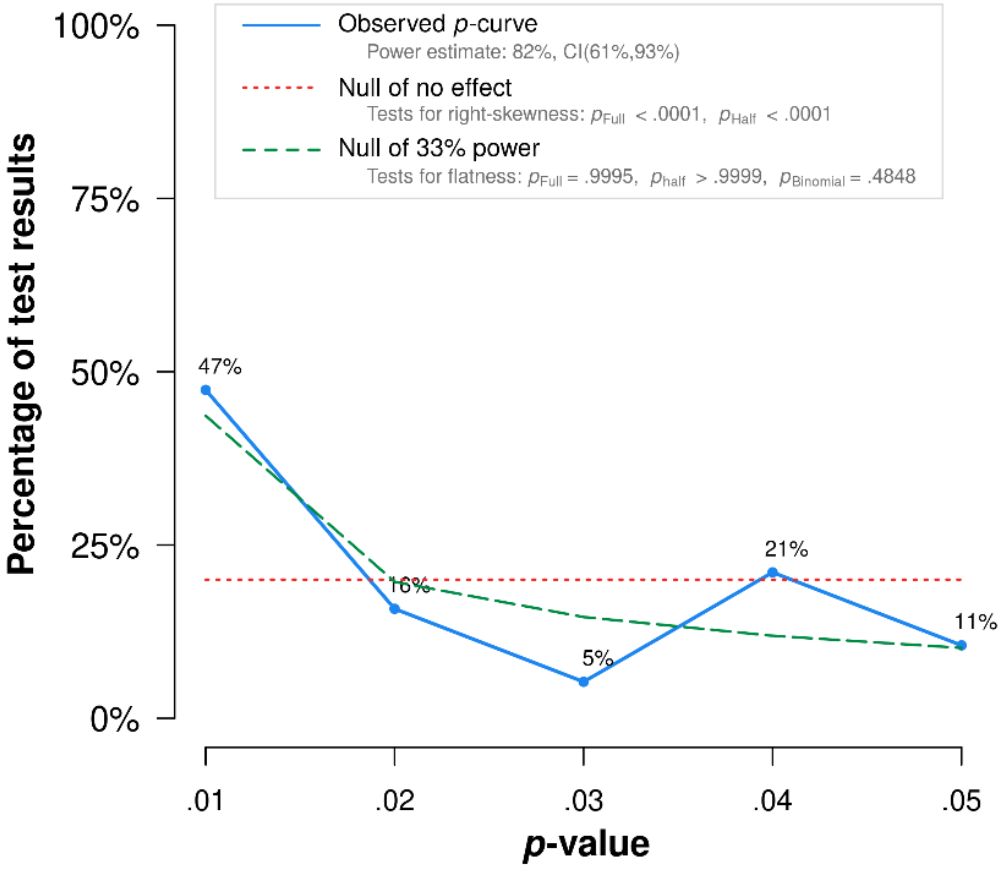
Figure F3. *Descriptive statistics and p-curve from intervention type Education.*
k = 50
n: min = 8, max = 78600, mean = 4679, sd = 14048, sum = 233592
WAAP: r = .03 (95% CI .02, .04), k = 13
Selection models: r = .21 (95% CI .04, .35)
p-curve analysis:



Note: The observed *p*-curve includes 27 statistically significant ($p < .05$) results, of which 19 are $p < .025$.
There were 23 additional results entered but excluded from *p*-curve because they were $p > .05$.

	Binomial Test *(Share of results p<.025)*	**Continuous Test** *(Aggregate with Stouffer Method)*	
		Full p-curve (p's<.05)	**Half p-curve (p's<.025)**
1) Studies contain evidential value. *(Right skew)*	*p*=.0261	*Z*=-14.79, *p*<.0001	*Z*=-18.21, *p*<.0001
2) Studies' evidential value, if any, is inadequate. *(Flatter than 33% power)*	*p*=.538	*Z*=9.79, *p*>.9999	*Z*=15.63, *p*>.9999
	Statistical Power		
Power of tests included in *p*-curve *(correcting for selective reporting)*	Estimate: 99% 90% Confidence interval: (99% , 99%)		

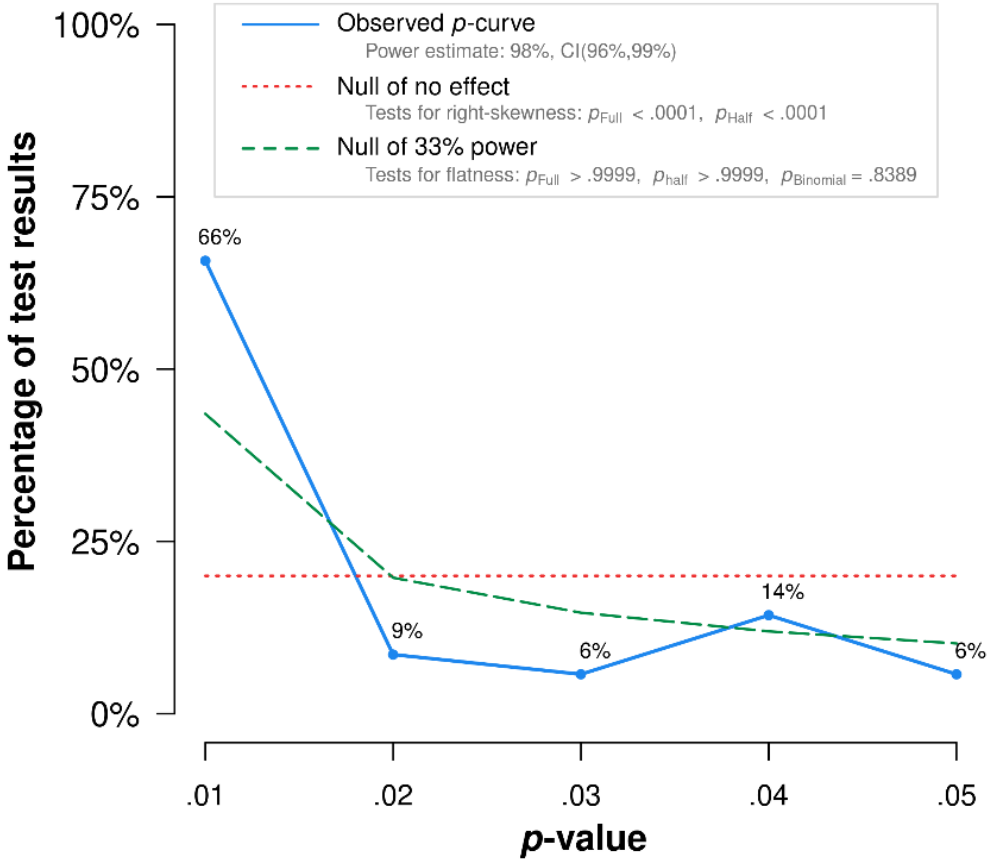
Figure F4. *Descriptive statistics and p-curve from intervention type Feedback*
k = 51
n: min = 8, max = 15000, mean = 1341, SD = 2988, sum = 68415
WAAP: r = .03 (95% CI .02, .04), k = 5
Selection models: r = .07 (95% CI .02, .11)
p-curve analysis:



Note: The observed *p*-curve includes 19 statistically significant ($p < .05$) results, of which 13 are $p < .025$.
There were 32 additional results entered but excluded from *p*-curve because they were $p > .05$.

	Binomial Test *(Share of results p<.025)*	**Continuous Test** *(Aggregate with Stouffer Method)*	
		Full p-curve (p's<.05)	**Half p-curve (p's<.025)**
1) Studies contain evidential value. *(Right skew)*	*p*=.0835	*Z*=-6.52, *p*<.0001	*Z*=-7.7, *p*<.0001
2) Studies' evidential value, if any, is inadequate. *(Flatter than 33% power)*	*p*=.4848	*Z*=3.28, *p*=.9995	*Z*=7.93, *p*>.9999
	Statistical Power		
Power of tests included in *p*-curve *(correcting for selective reporting)*	Estimate: 82% 90% Confidence interval: (61% , 93%)		

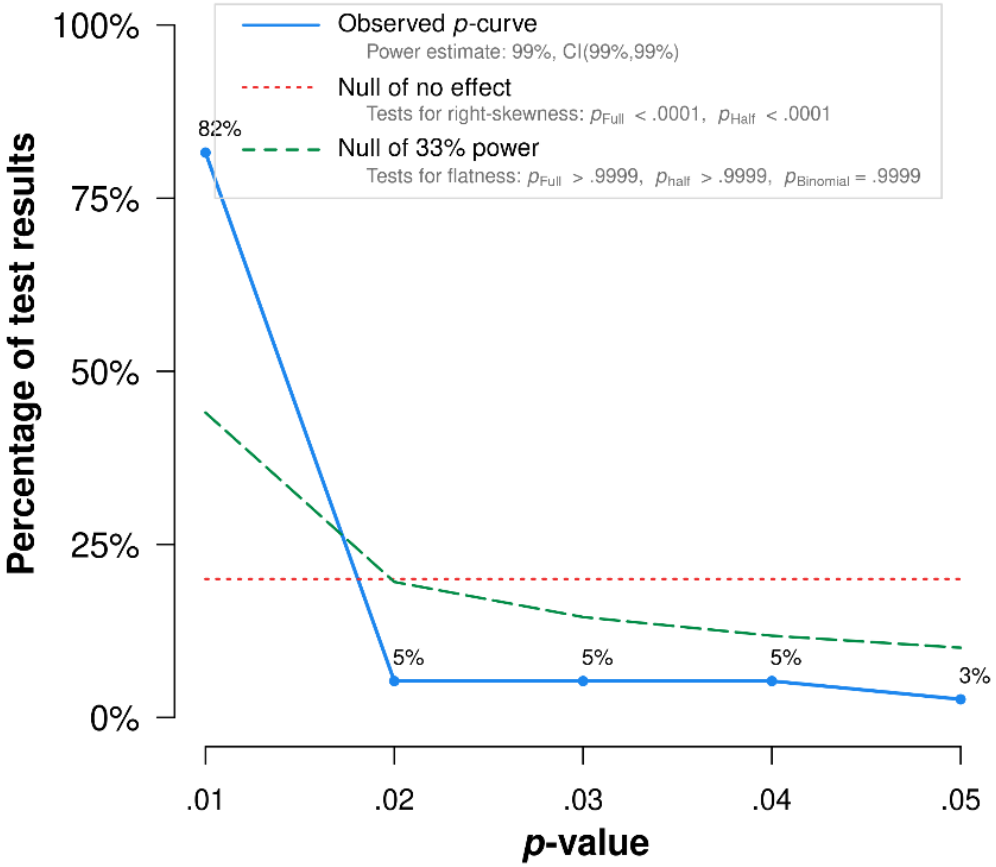
Figure F5. *Descriptive statistics and p-curve from intervention type Financial incentives.*
k = 75
n: min = 8, max = 15000, mean = 780, sd = 2326, sum = 58547
WAAP: r = .07 (95% CI .03, .11). k = 8
Selection models: r =.17 (95% CI .07, .27)
p-curve analysis:



Note: The observed *p*-curve includes 35 statistically significant ($p < .05$) results, of which 27 are $p < .025$.
There were 40 additional results entered but excluded from *p*-curve because they were $p > .05$.

	Binomial Test *(Share of results p<.025)*	**Continuous Test** *(Aggregate with Stouffer Method)*	
		Full p-curve (p's<.05)	**Half p-curve (p's<.025)**
1) Studies contain evidential value. *(Right skew)*	*p*=.0009	*Z*=-14.26, *p*<.0001	*Z*=-15.47, *p*<.0001
2) Studies' evidential value, if any, is inadequate. *(Flatter than 33% power)*	*p*=.8389	*Z*=9.03, *p*>.9999	*Z*=14.01, *p*>.9999
	Statistical Power		
Power of tests included in *p*-curve *(correcting for selective reporting)*	Estimate: 98% 90% Confidence interval: (96% , 99%)		

Figure F6. *Descriptive statistics and p-curve from intervention type Social comparison.*
k = 114
n: min = 6, max = 106467, mean = 5282, sd = 17897, sum = 602149
WAAP: r = .03 (95% CI .02, .04), k = 13
Selection models: r = .06 (95% CI -.02, .15)
p-curve analysis:



Note: The observed *p*-curve includes 76 statistically significant ($p < .05$) results, of which 67 are $p < .025$.
There were 38 additional results entered but excluded from *p*-curve because they were $p > .05$.

	Binomial Test *(Share of results p<.025)*	**Continuous Test** *(Aggregate with Stouffer Method)*	
		Full p-curve (p's<.05)	**Half p-curve (p's<.025)**
1) Studies contain evidential value. *(Right skew)*	*p*<.0001	*Z*=-32.5, *p*<.0001	*Z*=-33.21, *p*<.0001
2) Studies' evidential value, if any, is inadequate. *(Flatter than 33% power)*	*p*=.9999	*Z*=23.37, *p*>.9999	*Z*=28.56, *p*>.9999
	Statistical Power		
Power of tests included in *p*-curve *(correcting for selective reporting)*	Estimate: 99% 90% Confidence interval: (99% , 99%)		

Appendix F

Table C. *(1) number of meta-analyses; (2) number of meta-analytical effect sizes; (3) number of primary studies; (4) second-order, grand mean standardized difference estimate; (5) lower limit of the 95% confidence interval; (6) upper limit of the 95% confidence interval; (7) observed variance across first-order mean standardized difference estimates; (8) expected second-order sampling error variance; (9) estimated true variance across first-order mean standardized difference estimates (expected sampling error variance removed); (9) proportion of the variance across the first-order meta-analytic means that is due to second order sampling error variance.*

		Meta-analyses *m*	Meta-analytic effect sizes *n*	Primary studies *k*	Overall grand mean *d*	95% CI lower limit	95% CI upper limit	$S^2_{\bar{d}}$	$E\left(S^2_{e_{\bar{a}_i}}\right)$	$\hat{\sigma}^2_{\bar{d}}$
Overall										
	Averaged over meta-analyses	10	10	1178	0.310	0.302	0.320	0.158	< 0.001	0.158
	Averaged over subgroup ES	10	38	1041	0.303	0.292	0.313	0.175	< 0.001	0.175
	Unique ES	10	10	663	0.310	0.280	0.340	N/A	N/A	N/A
Interventions										
	Appeals	1	1	10	0.279	0.279	0.279	N/A	N/A	N/A
	Commitment	3	3	67	0.272	0.261	0.284	0.002	< 0.001	0.002
	Education	5	5	121	0.087	0.076	0.099	0.004	< 0.001	0.004
	Feedback	4	4	120	0.159	0.149	0.169	0.003	< 0.001	0.003
	Financial incentives	4	6	73	0.317	0.296	0.338	0.008	0.001	0.007
	Social Comparison	5	9	199	0.370	0.351	0.389	0.019	< 0.001	0.019
Outcomes										
	Conservation	6	13	404	0.254	0.239	0.270	0.026	< 0.001	0.026
	Consumption	2	3	18	0.197	0.178	0.217	0.002	0.002	< 0.001
	Littering	1	1	22	0.519	0.519	0.519	N/A	N/A	N/A
	Recycling	4	9	103	0.273	0.236	0.309	0.035	< 0.001	0.035
	Transportation	4	5	57	0.079	0.064	0.093	0.003	< 0.001	0.003