

Online Supplemental

Methods

Participant Exclusions

Table S1 shows our total sample size, exclusion size, final sample size, and demographics.

Study Name	Total Sample	Exclusion Size			Final Sample	Demographic		
		Attention	Error	Admit		M_{age}	SD_{age}	Female %
Pilot	97	13	21	24	49	28.5	10.8	39
Study 1	144	16	31	54	63	27.8	10.4	48
Study 2	196	20	64	96	63	24.7	6.8	60
Study 3	178	14	20	31	123	19.4	3.9	80
Study 4	113	14	14	34	62	18.6	1.3	68
Study 5	194	10	10	33	146	28.6	10.4	37
Study 6	198	18	47	46	105	26.5	8.9	44
Study 7	152	21	1	55	83	26.7	9.0	49
Study 8	199	26	1	58	124	26.2	9.4	63
Study 9	191	21	2	90	90	29.5	10.0	41
Study 10	298	22	n/a	120	166	27.5	9.9	46
Study 11	299	13	n/a	135	161	27.7	9.0	45
Study 12	52	10	0	4	39	19	1.24	69

Table S1. Total sample sizes, exclusion sizes, final sample size, and demographics. Attention

refers to our attention check exclusion, error refers to our performance check exclusion, and

admit refers to our self-admitted exclusion. For details, see https://aspredicted.org/KXV_6SH.

Individual Differences

Need For Cognition. The Need for Cognition Scale (Cacioppo et al., 1984) measures enjoyment of cognitively effortful activities in daily life. Example items include, “I prefer complex to simple problems,” “Thinking is not my idea of fun,” and “I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.” Response options were made on a scale ranging from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). Higher scores corresponded to a greater need for cognition.

Boredom Proneness. The Boredom Proneness Scale (Struk et al., 2016) measures propensity to be bored in daily life. Example items include, “I find it hard to entertain myself,”

“It takes more stimulation to get me going than most people,” and “In most situations, it is hard for me to find something to do or see to keep me interested.” Response options were made on a 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). Higher scores corresponded to higher boredom proneness.

Meaningfulness of Effort. The Meaningfulness of Effort Scale (Campbell, Chung, & Inzlicht, 2021) measures the meaningfulness of effort. Example items include, “When I push myself, what I’m doing feels important,” “Life would have no purpose if I never had to try,” and “When I exert myself, I feel connected to my ideal life.” Response options were made on a scale ranging from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). Higher scores corresponded to higher meaningfulness of effort.

Visual Attention Checks

In Study 4, every five trials, we asked participants “What was one digit that was displayed or typed by you?” If they chose the add task, they were required to type at least one number that they added. If they chose the do-nothing task, they were required to type one number that we flashed on the screen while they were doing nothing. They were given five seconds to respond. In Study 5, every ten trials participants were presented with the question, “Please select this, so we know that you are paying attention.” They were given two seconds to respond. We decided to not continue using these, because we did not want to increase additional demand for participants. We additionally included the full task load index, industriousness scale, and protestant work ethic scale in Studies 3 and 4; and the full task load index and industriousness scale for study 5; but decided not to continue using them to save study time.

Results

Median Choices

A Wilcoxon signed rank test confirmed that median choices show similar results as our parametric results (see Table S2).

Study	n	Median %	V	p
Study 1	63	60	1281	<.001
Study 2	63	60	1406	.003
Study 3	123	52	3916	.797
Study 4	62	53	923	.874
Study 5	146	48	3507	.017
Study 6	105	58	3223	.017
Study 7	83	53	1833	.056
Study 8	124	63	5493	<.001
Study 9	90	66	3368	<.001
Study 10	166	60	10114	<.001
Study 11	161	53	5918	.532
Study 12	39	40	151	.002

Table S2. The proportion of median demand choices in each study. Statistics are from a Wilcoxon signed rank one-sample test.

Meta-Analysis Choices

We also meta-analyzed effort choices across our studies without any exclusions and with only the self-admitted exclusion. While we preregistered these exclusion criteria, we re-analyze with different exclusion criteria to test the robustness of our results. Our meta-analysis revealed that, with only the self-admitted exclusion, people were indifferent between the two options, $d=.12$, 95% CI [-.08,.32], $z=1.15$, $p=.250$. Without any exclusions, people were indifferent between the demanding and do-nothing options, $d=-.02$, 95% CI [-.20, .16], $z=-.20$, $p=.842$. With only the attention check exclusion criteria, people were indifferent between the demanding and do-nothing options, $d=.05$, 95% CI [-.13, .23], $p=.558$. Forest plots are shown on Fig. S1. (self-exclusion), Fig. S2. (all exclusion), and Fig. S3. (attention check exclusions).

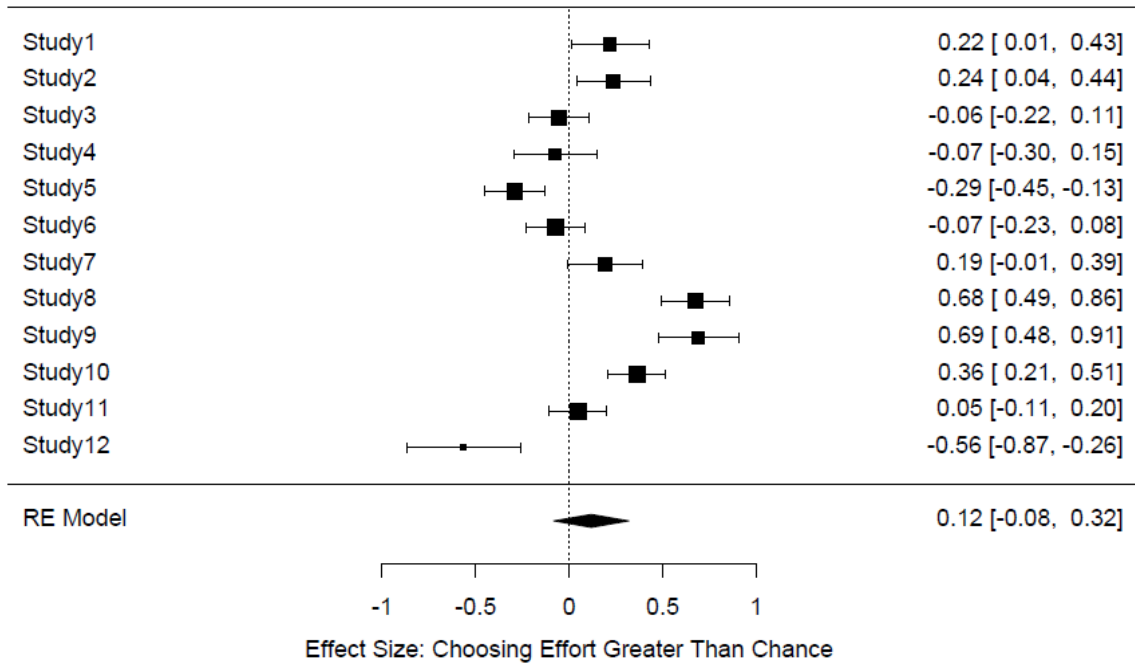


Fig. S1. Meta-analytic forest plot of effort choices across studies, with only self-admit exclusion.

Error bars represent 95% confidence intervals. Effect sizes represent Cohen's d from one-sample t -test comparing the proportion of effort choices with 0.5.

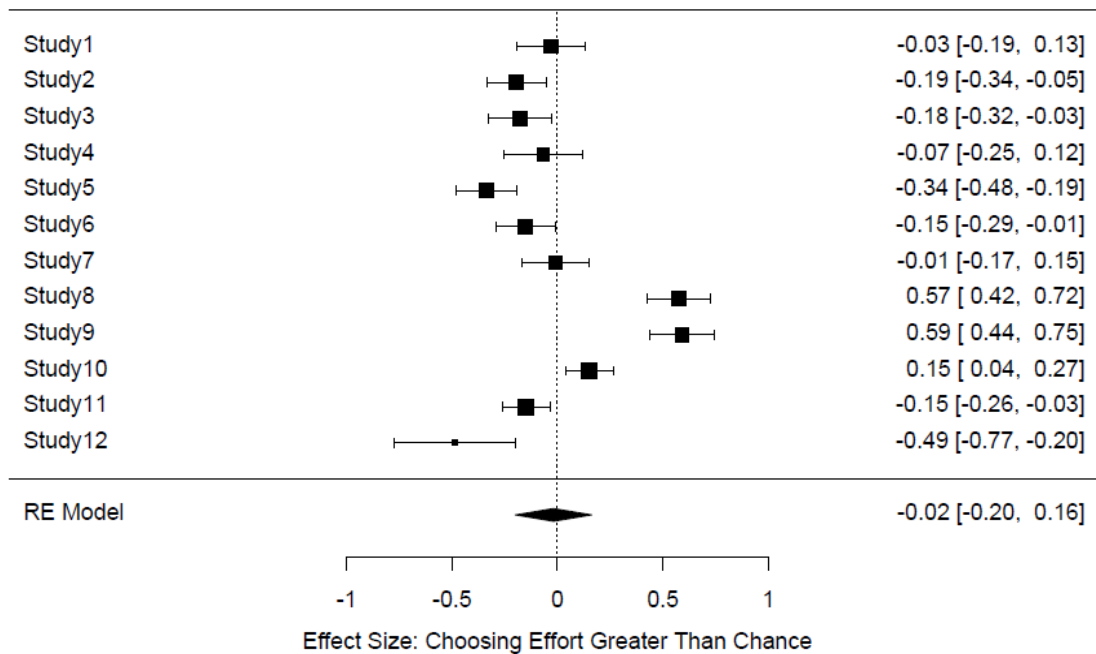


Fig. S2. Meta-analytic forest plot of demand vs. do-nothing choice across studies, with no exclusions. Error bars represent 95% confidence intervals. Effect sizes represent Cohen's *d* from one-sample *t*-test comparing the proportion of effort choices with 0.5.

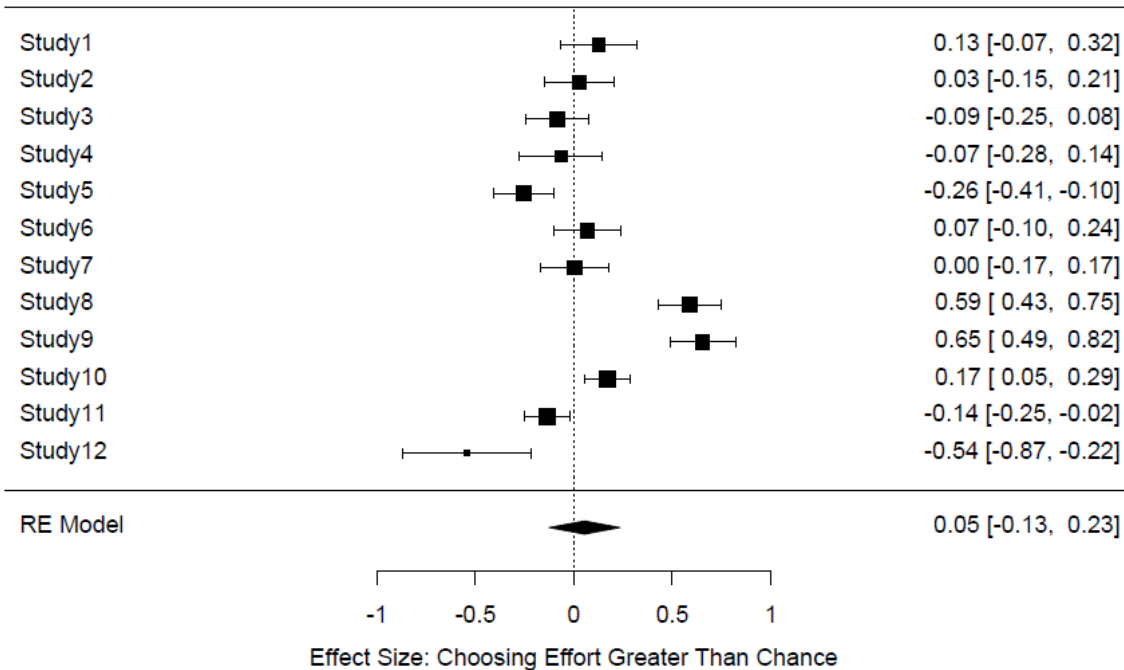


Fig. S3. Meta-analytic forest plot of demand vs. do-nothing choice across studies, with attention check exclusions. Error bars represent 95% confidence intervals. Effect sizes represent Cohen's *d* from one-sample *t*-test comparing the proportion of effort choices with 0.5.

Full Phenomenology Ratings

Study	Item	Do nothing Deck M	Do nothing Deck SD	Effort Deck M	Effort Deck SD	<i>d</i>	<i>p</i>
Study 1	How physically demanding?	.65	1.22	1.84	1.95	.51	<.001
Study 1	How hard did you have to work?	1.15	1.56	2.93	1.86	.87	<.001
Study 1	How insecure, discouraged, irritated, stressed, and annoyed?	1.46	1.78	2.48	2.06	.50	<.001

Study 1	How successful were you?	5.93	2.210	5.29	1.73	-.31	.001
Study 1	How hurried or rushed?	1.39	1.90	4.14	2.20	1.01	<.001
Study 2	How physically demanding?	.71	1.05	1.85	1.80	.73	<.001
Study 2	How hard did you have to work?	.98	1.27	3.30	1.88	1.12	<.001
Study 2	How insecure, discouraged, irritated, stressed, and annoyed?	1.43	1.64	2.57	2.26	.44	.001
Study 2	How successful were you?	5.86	2.10	4.89	1.75	-.41	.003
Study 2	How hurried or rushed?	1.48	1.69	4.20	2.23	1.03	<.001
Study 3	How physically demanding?	.75	1.33	1.89	1.85	.62	<.001
Study 3	How hard did you have to work?	1.01	1.69	3.21	1.82	1.05	<.001
Study 3	How insecure, discouraged, irritated, stressed, and annoyed?	1.83	2.13	1.98	1.85	.05	.537
Study 3	How successful were you?	5.92	2.10	5.49	1.68	-.20	.018
Study 3	How hurried or rushed?	1.17	1.79	3.69	2.25	.93	<.001
Study 10	How physically demanding?	.47	.97	1.40	1.60	.7	<.001
Study 10	How hard did you have to work?	.56	1.10	2.16	1.84	1.05	<.001
Study 10	How insecure, discouraged, irritated, stressed, and annoyed?	.95	1.47	1.15	1.44	.13	.136
Study 10	How successful were you?	5.79	2.34	5.92	1.45	.07	.512
Study 10	How hurried or rushed?	.87	1.34	1.57	1.38	.52	<.001
Study 10	How much did you think about how sad or upset you felt?	.85	1.31	.90	1.25	.04	.685
Study 10	How meaningful?	1.07	1.54	2.79	2.02	.96	<.001

Study 11	How physically demanding?	.40	.98	1.31	1.63	.68	<.001
Study 11	How hard did you have to work?	.57	1.29	2.29	1.84	1.08	<.001
Study 11	How insecure, discouraged, irritated, stressed, and annoyed?	.89	1.42	1.08	1.39	.14	.095
Study 11	How successful were you?	6.38	1.66	6.16	1.32	-.15	.135
Study 11	How hurried or rushed?	.91	1.51	1.63	1.44	.49	<.001
Study 11	How much did you think about how sad or upset you felt?	1.23	1.66	.87	1.3	-.24	.004
Study 11	How meaningful?	1.31	1.84	2.71	2.07	.72	<.001
Study 12	How physically demanding?	.93	1.29	1.20	1.36	.26	.115
Study 12	How hard did you have to work?	.99	1.15	1.74	1.43	.68	<.001
Study 12	How insecure, discouraged, irritated, stressed, and annoyed?	1.46	1.66	1.52	1.57	.06	.715
Study 12	How successful were you?	6.15	.96	6.29	1.01	.19	.251
Study 12	How hurried or rushed?	1.02	1.23	2.09	1.78	.68	<.001
Study 12	How much did you think about how sad or upset you felt?	.88	1.2	.78	.99	-.12	.467

Table S3. Descriptive statistics of phenomenology across studies. Effect sizes represent Cohen's *d* from paired t-tests.

Social Desirability

The social desirability scale (He et al., 2015) measures the tendency for participants to respond in a manner that will be viewed favorable by others. Example items include, "I am satisfied when I get my way", "I am irritated by people who ask favors", and "I do things my way." Response options were made on a scale ranging (from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*)). Higher

scores corresponded to greater social desirability. We included this scale in Studies 10-12. Table S4. shows the association between social desirability and effort choices.

Study	Mean	SD	<i>r</i>
Study 10	3.83	0.46	.09 [-.07, .24]
Study 11	3.81	0.43	.13 [-.03, .28]
Study 12	3.93	0.37	.00 [-.32, .32]

Table S4. Social desirability means and standard deviations. Effect size corresponds to correlation between social desirability and effort choices.