Fear of the Light?
Transparency does not reduce the effectiveness of nudges. A data-driven review

Hendrik Bruns, Joint Research Centre of the EC, Belgium

Adrien Fillon, SInnoPSis, Dept. of Economics, UCY

Zacharias Maniadis, SInnoPSis, Dept. of Economics, UCY and University of Southampton

Yavor Paunov, Royal Institute of Technology, Stockholm

Nudge and its discontents

- Nudges systematically change behaviour by affecting choice architecture without using financial incentives or significant restrictions.
  - (Thaler and Sunstein, 2008)

- Use of nudges has been one of the most impressive developments in public policy in the last 15 years.
  - (Hallsworth, 2023)

- Bovens (2009) argue that nudges are more ethically permissible if:
  - They target behaviours that individuals themselves would rate as desirable
  - They do not disturb individuals' feelings of internal consistency
  - They allow people to maintain control over their behaviour

Transparency

- Covert nudges may limit autonomous choice and can be perceived as external attempts to drive people's behaviour without their consent.
  - Typically manipulating automatic and non-deliberative processes

- Nudge designers may enhance transparency over what nudge is attempting to do (and via which means) and protect agency.

- Calls for a minimum level of transparency when these interventions are used.
  - (House of Lords, 2011)

What does theory predict? Does transparency undermine the effects of nudges?

- People strive for self-determination.
  - (Deci and Ryan, 1985)
- Can perceive nudging as a limitation to their freedom of choice and deliberately ignore or resist a nudge to reinstall that freedom.
  - (Brehm, 1966, Paunov et al., 2019)
- Conditioning literature suggests that people may respond more favourably when they are aware of the nudging intervention.
  - (Gorn, Jacobs, and Mana, 1987)
- ‘Third person effect’: participants do not view themselves prone to manipulation, potentially moderating the possible effect of reactance.
  - (Perloff, 1993)

What do we know?

- *Informal* surveys have considered the effects of transparency on efficacy of nudges. Transparency does not seem to reduce nudge effects.
  - Michaelsen and Sunstein (2023), DeRidder, Kroese and van Gestel (2022), Marchiori, Adriaanse and de Ridder (2016)

- Sunstein (2016) provides evidence that the public does not believe that transparency reduces the effectiveness of nudges.

- Michaelsen and Sunstein (2023) made strong claim: "… *when a disclosure is presented along with a default nudge, the effect on behavior does not diminish*. To that extent, Boven's conjecture has been falsified." (p. 37).

- Michaelsen (2023) draws conclusions about the limitations of the current literature.

Our Contribution: a Formal Meta-Analysis

- Informal reviews prone to risks of incomplete, selective, and subjective coverage of literature.
- Missing rigorous quantitative analysis that can help us summarize limitations in generalizing and draw inferences from the current literature.
- Such an analysis of limitations can also guide future research and suggest best intervention practices.
- Take stock in systematic and rigorous manner of whether coupling nudges with a transparency message – thus alleviating most ethical concerns – would undermine their efficacy.

Methods

- Searched using Scopus and the Web of Science Core Collection.

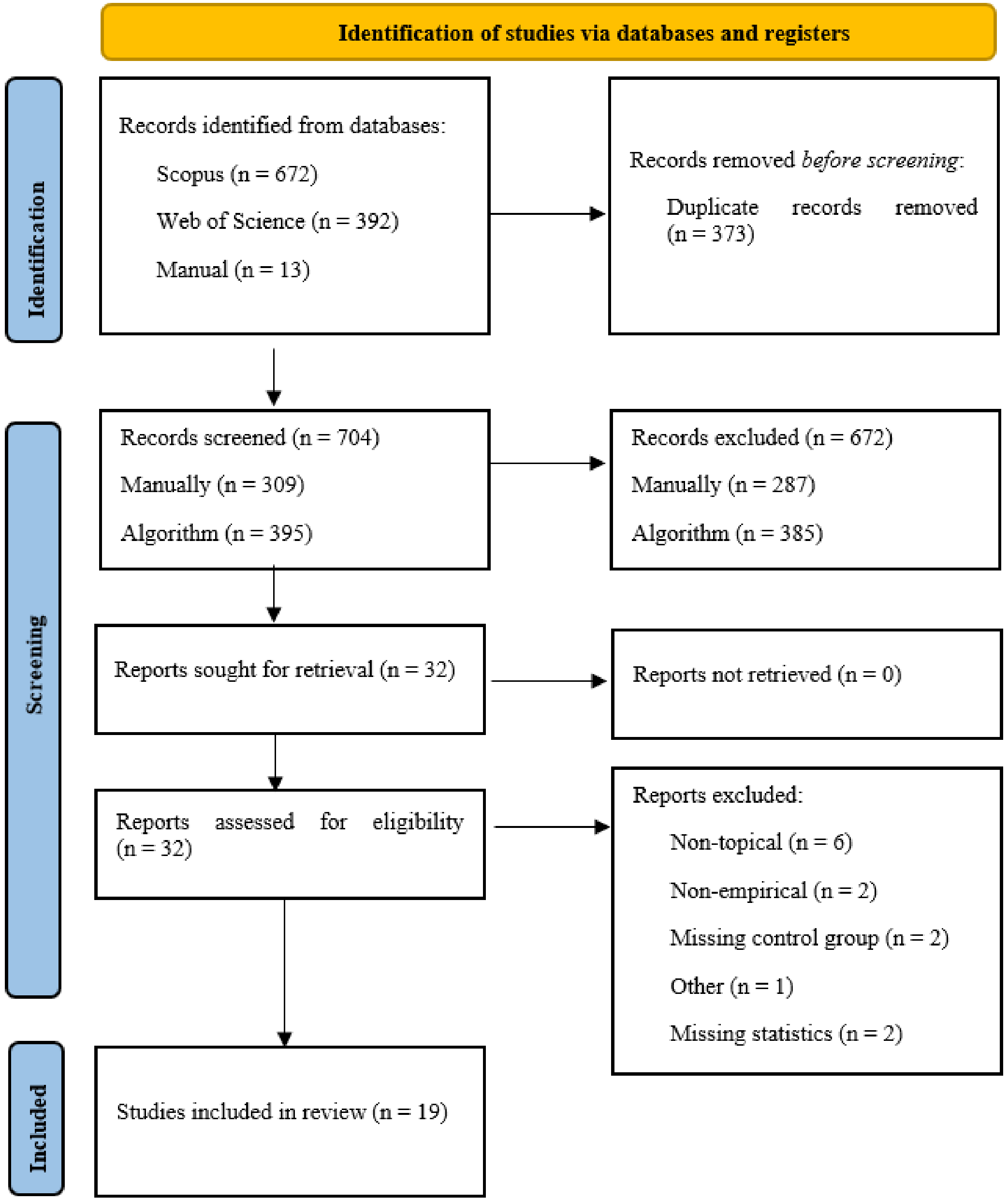
- After excluding duplicates based on the Digital Object Identifier (DOI), we screened titles and abstracts using ASReview (https://asreview.nl/).

- Screened all articles manually and independently. After initial screening of titles and abstracts, full texts were reviewed.

- Looked at other articles by identified authors in the field, to check whether there are relevant papers that we may have missed.

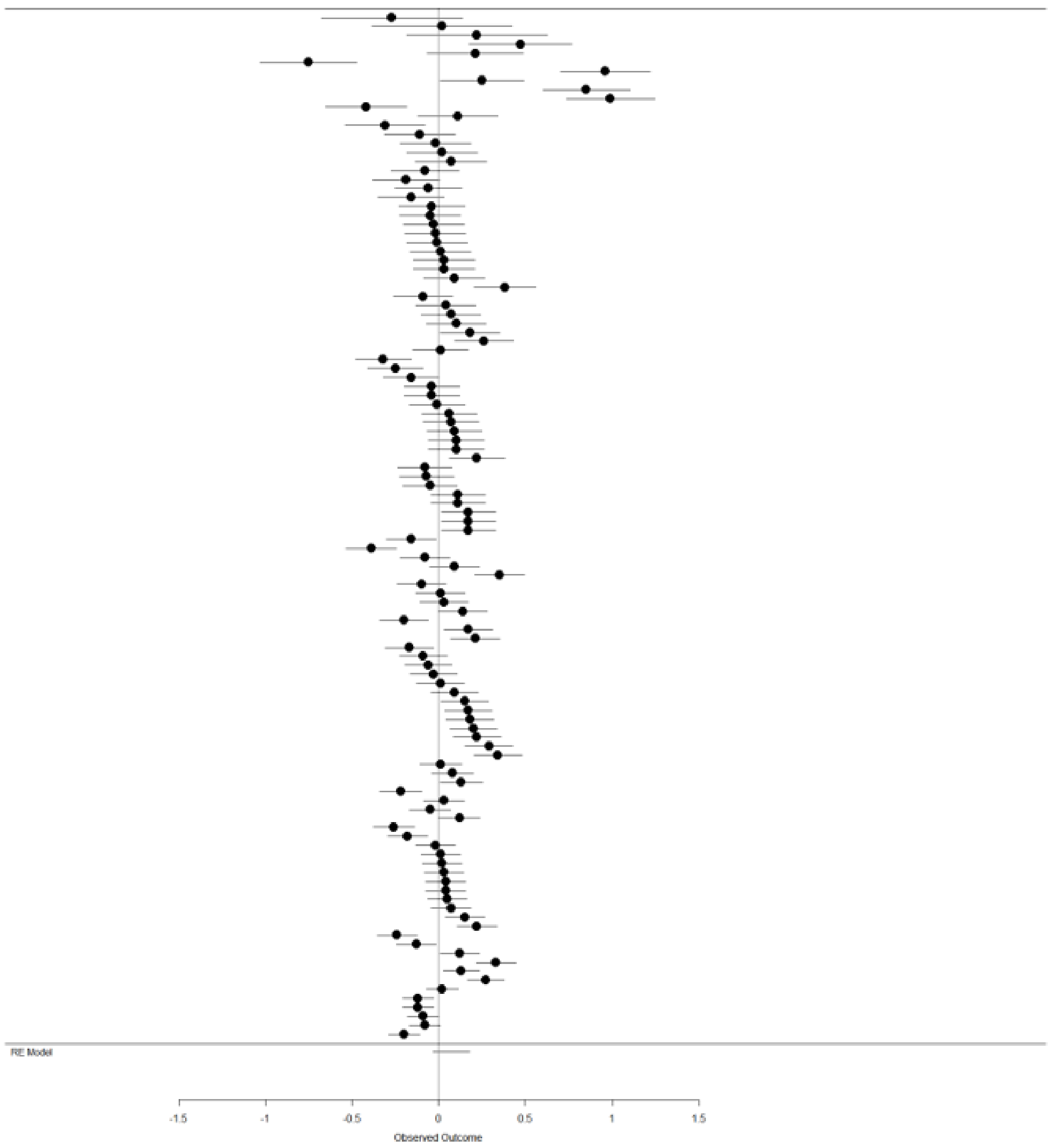
- For all the articles, titles, abstracts, tables, and methods sections were scanned to identify the relevance of a given source.

Figure 1: *PRISMA Flowchart of included studies*

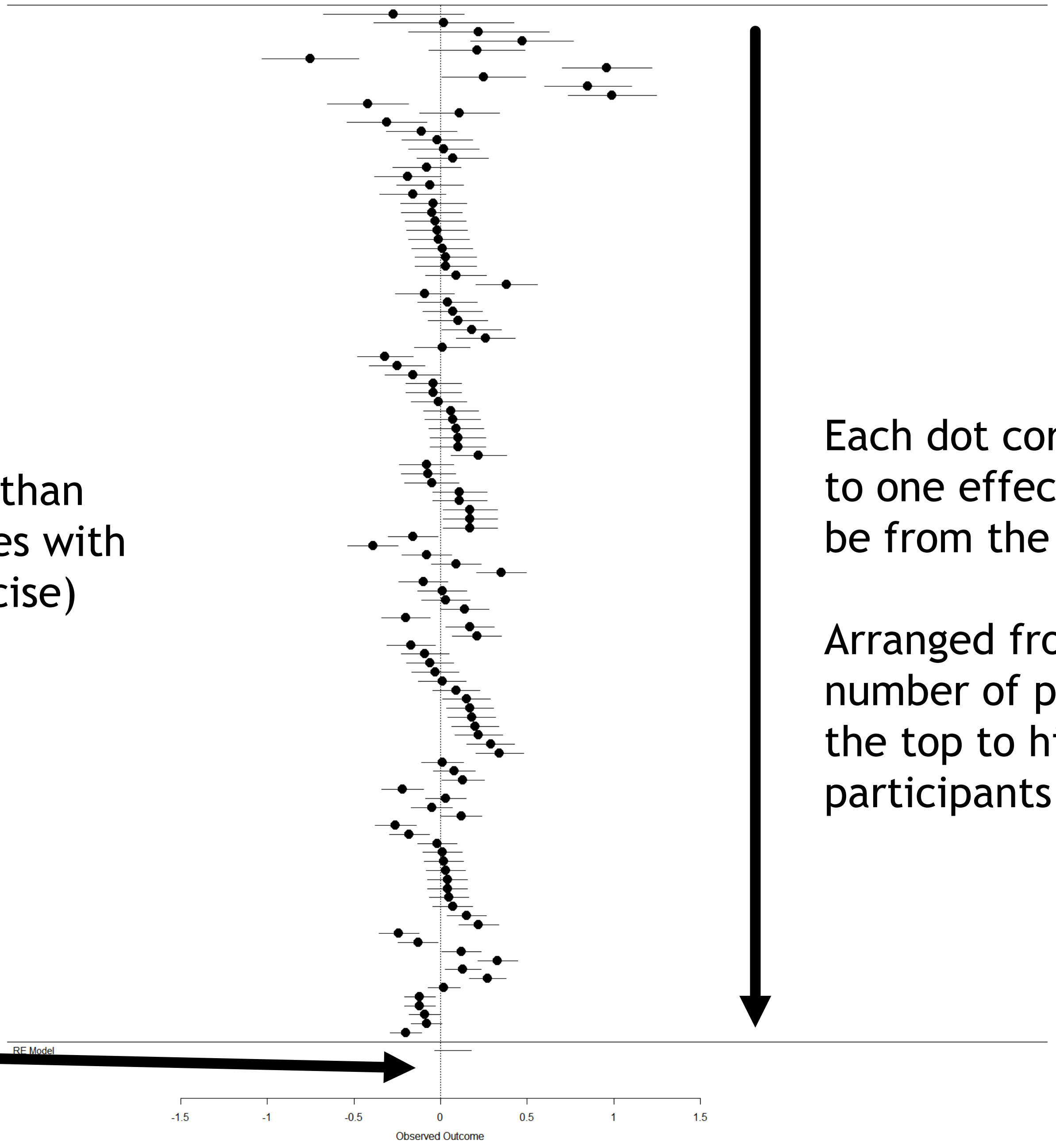


Results

Figure 2: *Forest plots of all effect sizes (n = 19, k = 114) included in the meta-analysis with their corresponding 95% confidence intervals. Effect sizes are arranged from the smallest to the highest sample size.*



This is called a Caterpillar Plot



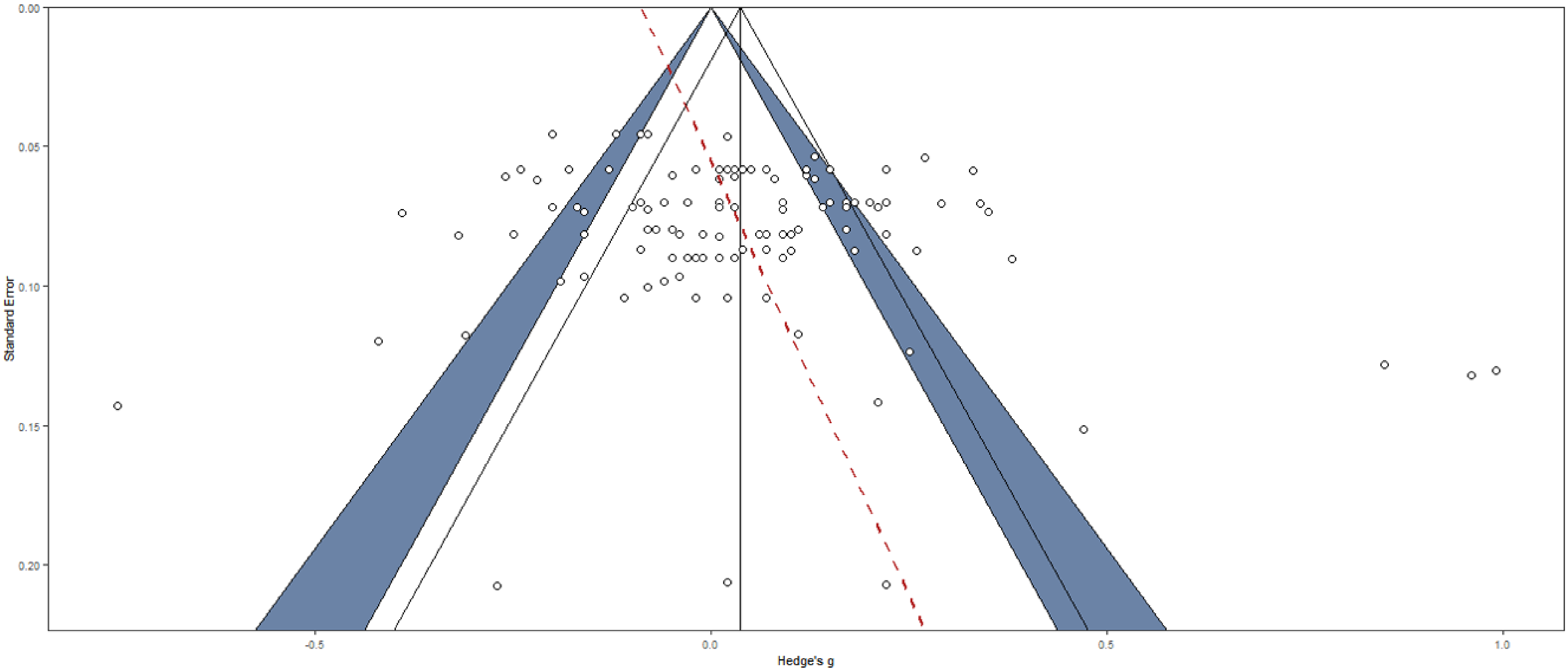
Meta-analytic estimate

More heterogeneity on the top than bottom, which is normal (studies with fewer participants are less precise)

Each dot corresponds to one effect size but can be from the same study

Arranged from the lower number of participants at the top to higher number of participants at the bottom

Figure 3: *Funnel plot displaying each observation as a function of its effect size and SE.*

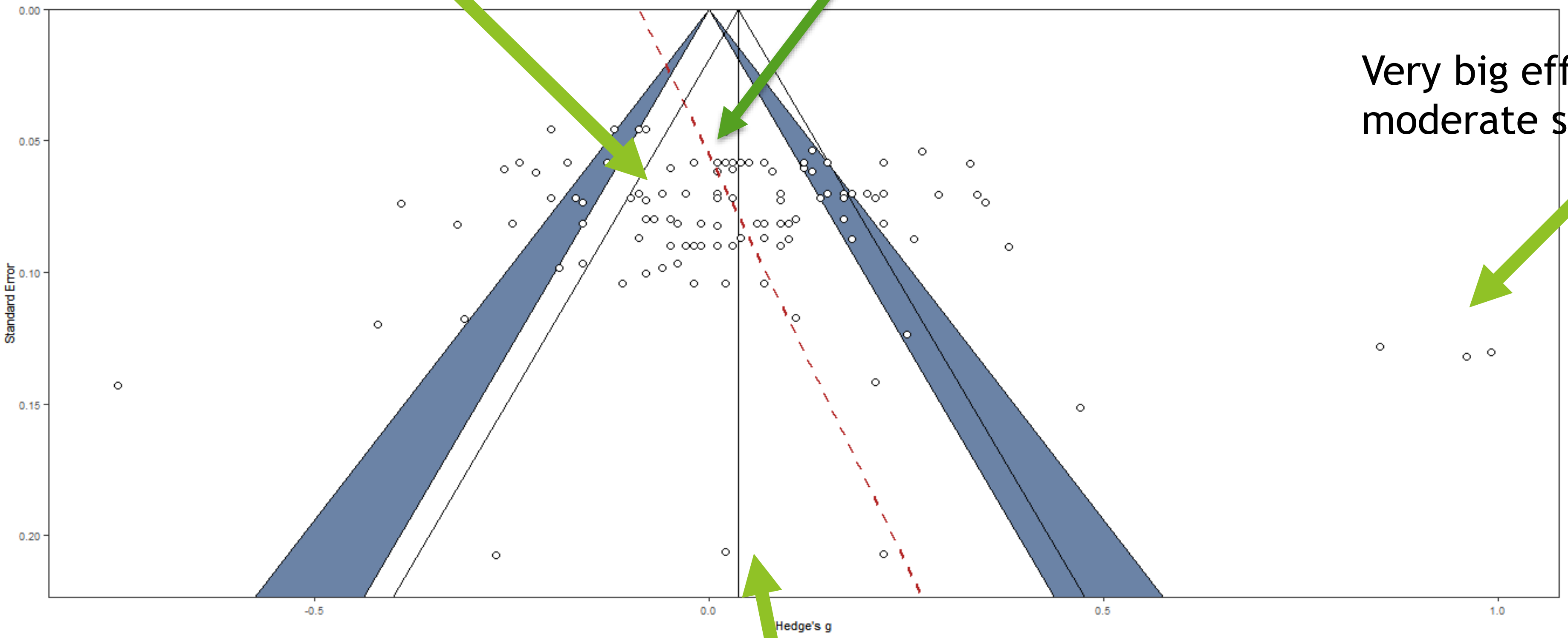


X axis is the effect size, Y axis is standard errors. If a true effect exists, the dots should form a pyramid around it (top of pyramid is where standard errors= 0). Dont see pyramid at all, with lots of effect sizes outside the blue pyramids. Possibility of publication bias.

Red line is the egger regression test, should be vertical.
Here it is not, indicating that the 3 points on the right side are weird

Not in a shape of pyramid

Very big effect size,
moderate standard errors



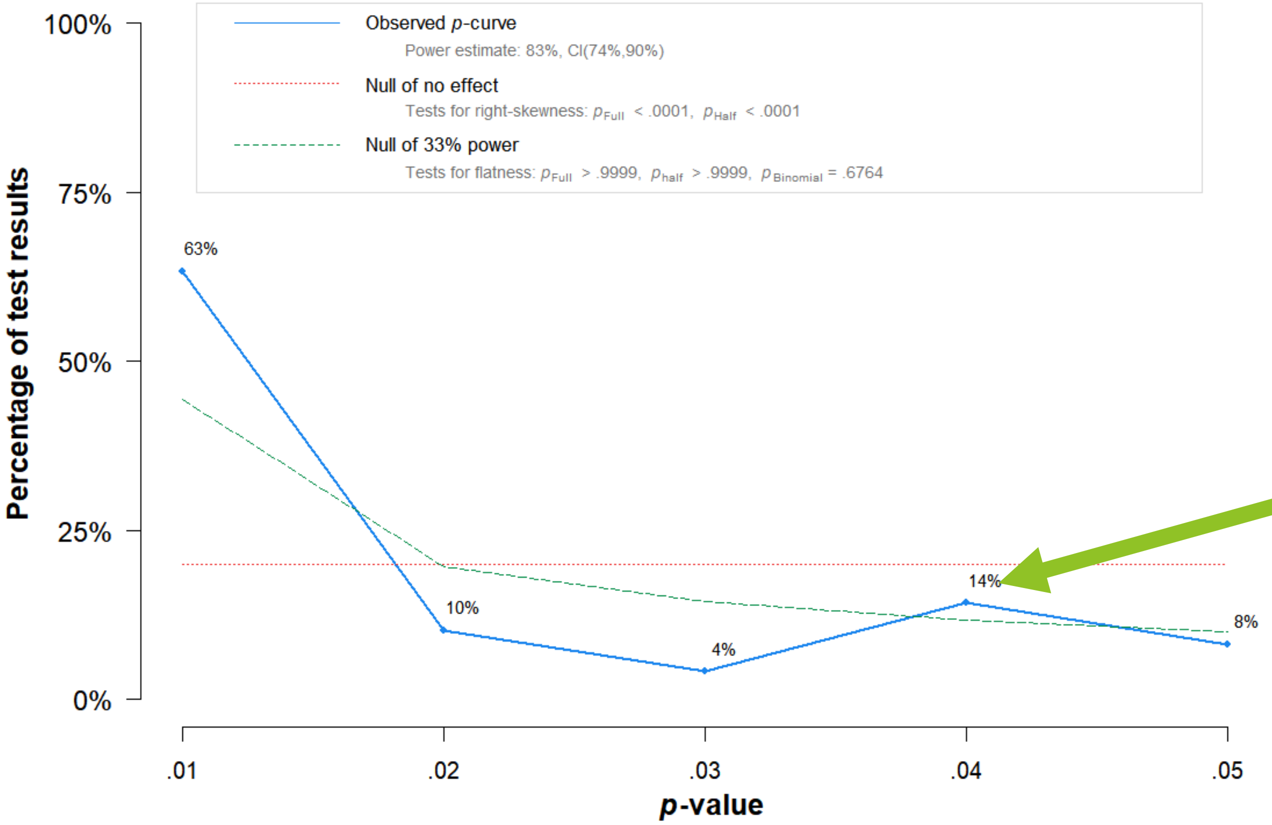
Very small effect size,
big standard errors

Table 2: *Moderator analysis*

Effect	*k*	*g*	95% CI	Test statistic	*P*
Type of Experiment				$F(2, 108) = 0.03$	0.96
Online	99	0.08	[-0.06, 0.21]		
Laboratory	10	0.06	[-0.02, 0.14]		
Field	2	0.02	[-0.07, 0.12]		
Intervention Category				$F(1, 112) = 0.14$	0.71
Decision Structure	105	0.10	[-0.02, 0.22]		
Decision Information	8	-0.06	[-0.21, 0.09]		
Decision Assistance	1	-0.05			
Domain				$F(1, 112) = 0.01$	0.92
Environment	43	0.04	[-0.01, 0.10]		
Food	2	0.02	[-0.07, 0.12]		
Health	14	-0.18	[-0.43, 0.07]		
Other	27	0.17	[-0.10, 0.43]		
Pro-social	28	0.03	[-0.02, 0.08]		

P-curve analysis. You take all the significant p-values from the dataset.

You should have most of them at .01 and a uniform distribution between .02 and .05 to not indicate p-hacking.



Note: The observed *p*-curve includes 49 statistically significant ($p < .05$) results, of which 36 are $p < .025$.
There were 65 additional results entered but excluded from *p*-curve because they were $p > .05$.

Small increase at .04 but not above the red line
No indication of p-hacking.

Conclusions

- Empirical literature does not suggest that transparency affect efficacy of nudges.
- However, major limitations in the literature:
  - Low probability that publication bias might have driven the results, but number of studies is limited and dominated by a few research teams.
  - No meaningful meta-analytic moderators that could explain in which settings transparency can affect the efficacy of nudges.
  - Most studies focused on default nudges in an online setting with weak incentives, and only two were conducted in a field setting with an actual meaningful nudge.
  - Majority of studies used a default choice.

- Overall, literature supports the view that nudges need not operate ‘in the dark’.