

The (boring) methods behind nudging

Erik Gahner Larsen

(@erikgahner / egl.dk)

Copenhagen Behavioural Economics Network
April, 2016

Me

- ▶ Erik Gahner Larsen
- ▶ Department of Political Science, University of Southern Denmark
- ▶ Political science, public policy, causal inference
- ▶ **Not** an expert on nudging
 - ▶ Outsider
 - ▶ Positive, optimistic

What's up?

- ▶ Everybody wants to nudge the world
 - ▶ Behavioral economics (economists, psychologists)
 - ▶ Social sciences (law, sociology, political science)
 - ▶ Business (communicators, designers, advertisers)
 - ▶ Government employees (civil servants)
 - ▶ etc.
- ▶ Nudging is popular: for good reasons
 - ▶ But not the only game in town

Nudging in context

- ▶ Nudging: one *distinct* mode of governance among others (Mols et al. 2015)
 - ▶ hierarchy
 - ▶ markets
 - ▶ networks
 - ▶ persuasion
 - ▶ *nudging*
- ▶ To nudge or not to nudge ...
 - ▶ ... is not the only question
 - ▶ Multiple different nudges

Why is nudging so popular?
Why are we all here?

Nudging and science

Nudging	Science
Innovative	Old-fashioned
Revolutionary	Incremental
Exciting	Boring
Counterintuitive	Intuitive

So, is nudging unscientific?

So, is nudging unscientific?

No

So, is nudging unscientific?

Sometimes

So, is nudging unscientific?

Nudging = Science + bullshit

When is nudging scientific?

- ▶ Isolating science from bullshit: methods
- ▶ Important: theoretical criteria (theory, concepts)
- ▶ Again: a lot of people create nudges
- ▶ Why methods?
 - ▶ Choice architects are also human
 - ▶ Cognitive biases and limitations are universal
- ▶ We need specific *methodological* criteria to evaluate nudges

Three methodological criteria

- ▶ Three questions we have to ask
- ▶ Relate to different (validity) dimensions
- ▶ Methodological criteria
 - ▶ Causality (*internal validity*)
 - ▶ Relevance (*ecological validity*)
 - ▶ Generalizability (*external validity*)

Criteria 1: Causality

- ▶ How sure can we be that a nudge is actually driving a behavioral change?
- ▶ Solution: Experiments
 - ▶ Control and stimuli groups
 - ▶ Placebo tests/groups
 - ▶ Random assignment
 - ▶ Double blind assignment procedure
- ▶ (Honest) question: How often do we see such strong designs?

Criteria 2: Relevance

- ▶ How sure can we be that a nudge is actually implemented in a relevant setting?
- ▶ Solution: *Field* experiments
 - ▶ Experimental design (randomization and whatnot)
 - ▶ Real-world settings
 - ▶ Relevant changes in the environment
 - ▶ Theoretically motivated
 - ▶ Well-defined
 - ▶ Realistic

Criteria 3: Generalizability

- ▶ How sure can we be that a nudge working in one context works in another?
- ▶ Solution: *Multiple* field experiments
 - ▶ Experimental design (randomization and whatnot)
 - ▶ Multiple real-world settings
 - ▶ *Different* relations to other modes of governance
- ▶ How often do we see nudges implemented (= evaluated) in different contexts?

Big problems, big nudges?

- ▶ Nudging is the answer, but to *all* questions?
- ▶ Tom Goodwin (2012): "[Nudging is] not an effective strategy for bringing about the kind of behavioural changes required to solve society's 'big problems' – problems around climate change and public health, for example."

What about economic incentives?

- ▶ Incentives help people ...
 - ▶ exercise (Charness and Gneezy 2009)
 - ▶ stop smoking (Volpp et al. 2009)
 - ▶ lose weight (John et al. 2011)
 - ▶ ... but can also crowd out intrinsic motivations (Fehr and Falk 2002, Gneezy et al. 2011)
-
- ▶ When is a nudge working *better* than a financial incentive?
 - ▶ How often do we test nudges *vis-a-vis* other modes of governance (e.g. incentives)?
 - ▶ From a policy-maker perspective, that is (often) the relevant counterfactual

Example: Two nudges for number one

- ▶ Case/problem: Public urination at the Copenhagen Central Station
- ▶ **Honest** question: How do we solve this problem?

Example: Two nudges for number one

- ▶ Solution/nudge: Stickers (showing the direction to the toilet)

- ▶ Criteria 1 (causality): How to test this causally?
- ▶ Criteria 2 (relevance): How relevant is this test?
- ▶ Criteria 3 (generalizability): How will this nudge work across contexts?

Example: Two nudges for number one

- ▶ Causality in practice
 - ▶ No control group
 - ▶ Pre-test post-test measurements
 - ▶ Different groups, different points in time
 - ▶ Different "nudges"
 - ▶ Stickers
 - ▶ Container
 - ▶ Results?
 - ▶ 500 people urinated in the week before the nudge, 500 in the week after
 - ▶ Interpretation? The "experiment" was a huge success

Example: Two nudges for number one

- ▶ Relevance
 - ▶ Is it feasible?
 - ▶ Containers, seriously?
 - ▶ Do we get the desired goal?
 - ▶ Short-term vs. long-term goals
 - ▶ What about other nudges?
 - ▶ What about other modes of governance?
 - ▶ ↑ fine (not using toilet)?
 - ▶ ↓ cost (using toilet)?

Example: Two nudges for number one

- ▶ Generalizability
- ▶ Different contexts
 - ▶ People
 - ▶ Environments
 - ▶ Stimuli (*nudges*)
 - ▶ Time

Conclusion

- ▶ Method matter; apply the three criteria every time you encounter a nudge
 - ▶ Does it provide *causal* evidence? (e.g. randomized trials, experiments)
 - ▶ Does it provide a relevant test? (e.g. alternative modes of governance)
 - ▶ Does it provide arguments for why it should work in another context? (e.g. replications, a strong theory)
- ▶ Look for quality cues (on average, you are better off)
- ▶ Useful heuristics
 - ▶ Incentives
 - ▶ Peer-reviewed
 - ▶ Replications
 - ▶ Experiments
 - ▶ Academic affiliation

Thanks :-)