

Ethics of influence

Leading Principles Underlying Responsible Nudging

Christine Clavien
christine.clavien@unige.ch

Simple frictions (that make it a bit more difficult to put suicidal thoughts into practice) can dramatically affect suicide rate

In the early sixties the source of the gas in people's ovens began to change following the discovery of North Sea gas off the coast of Britain. Unlike the previous gas that had been derived from coal, the new North Sea 'natural gas' had much lower levels of carbon monoxide.

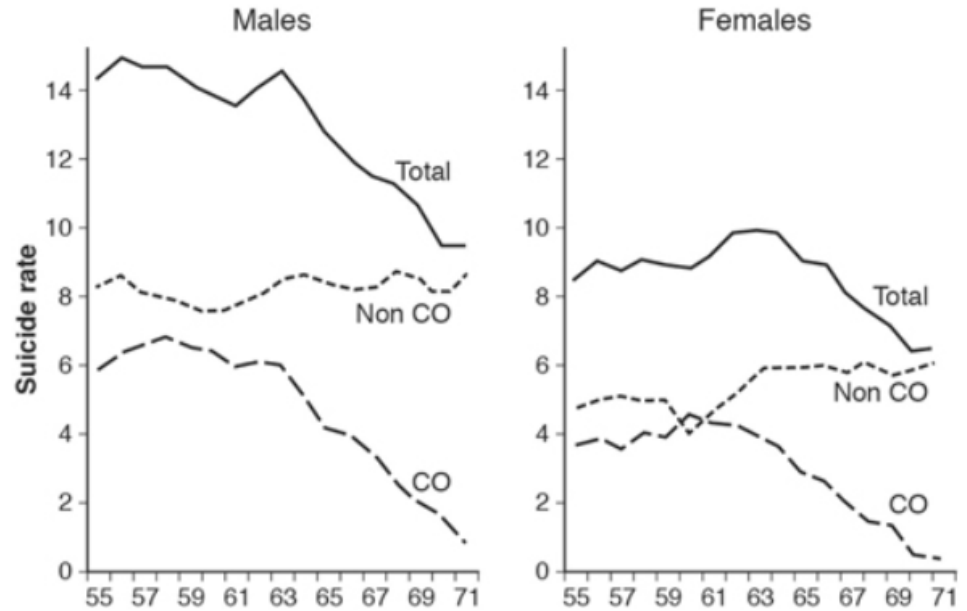


Figure 9. Trends in suicide rates per 100,000 in England and Wales, 1955–71, involving carbon monoxide and not involving carbon monoxide.

1,112 parole board hearings by eight judges, over ten months

!!! Some confounding factors may explain part of the results. But still, the general picture remains

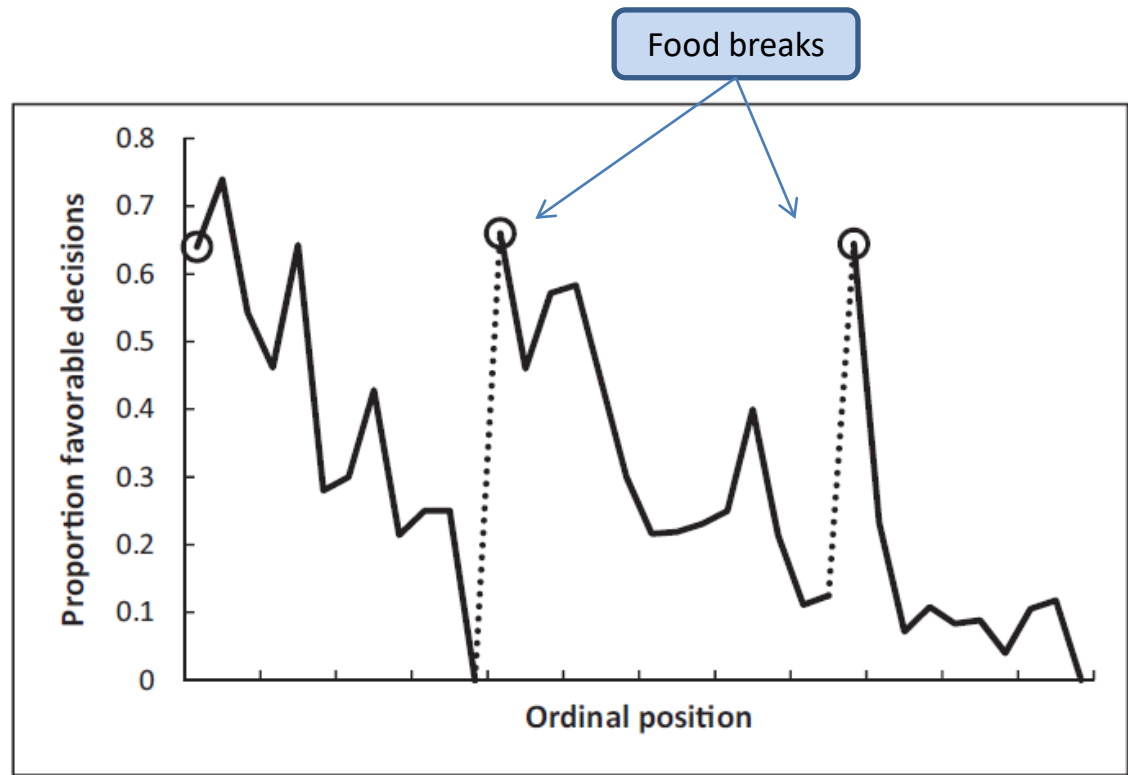
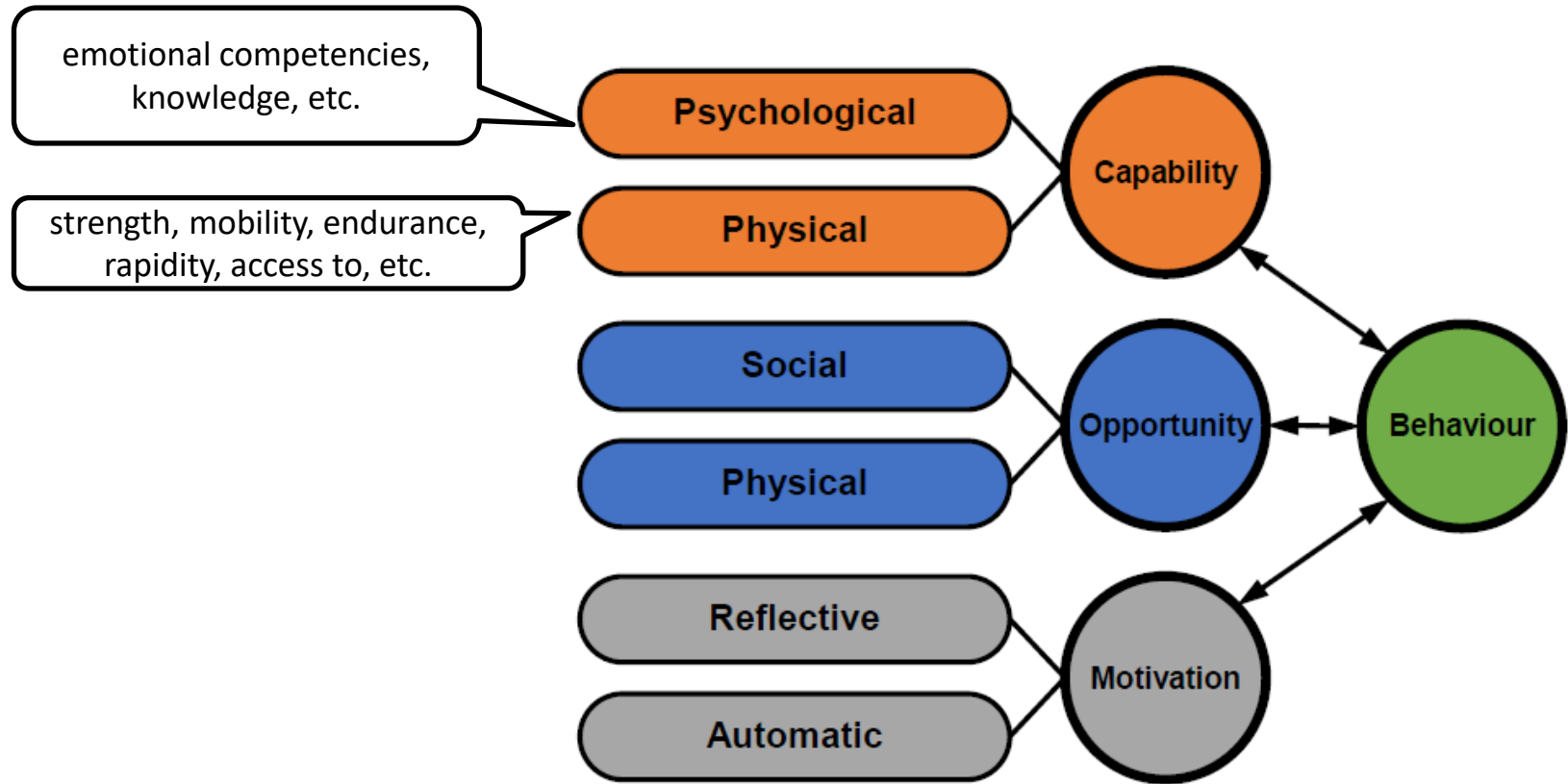
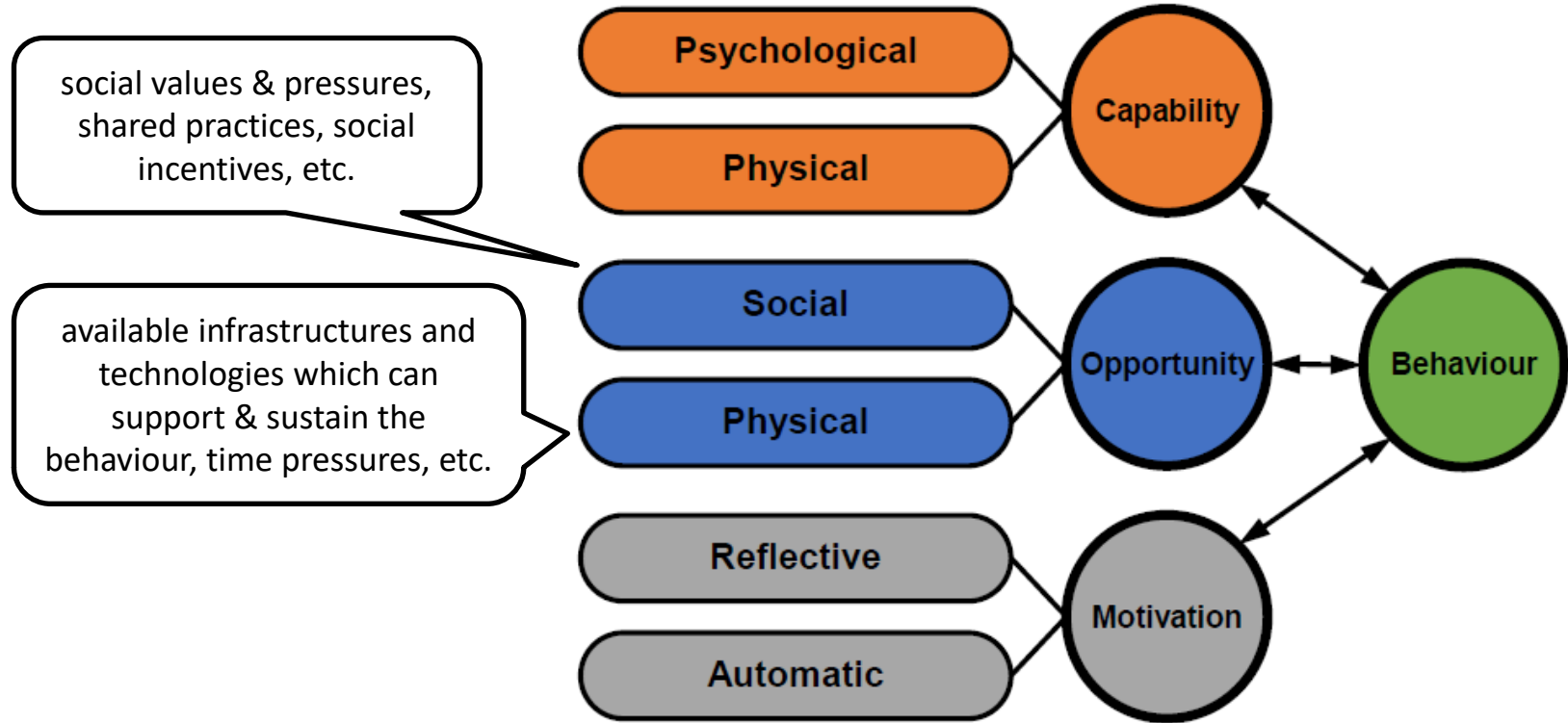
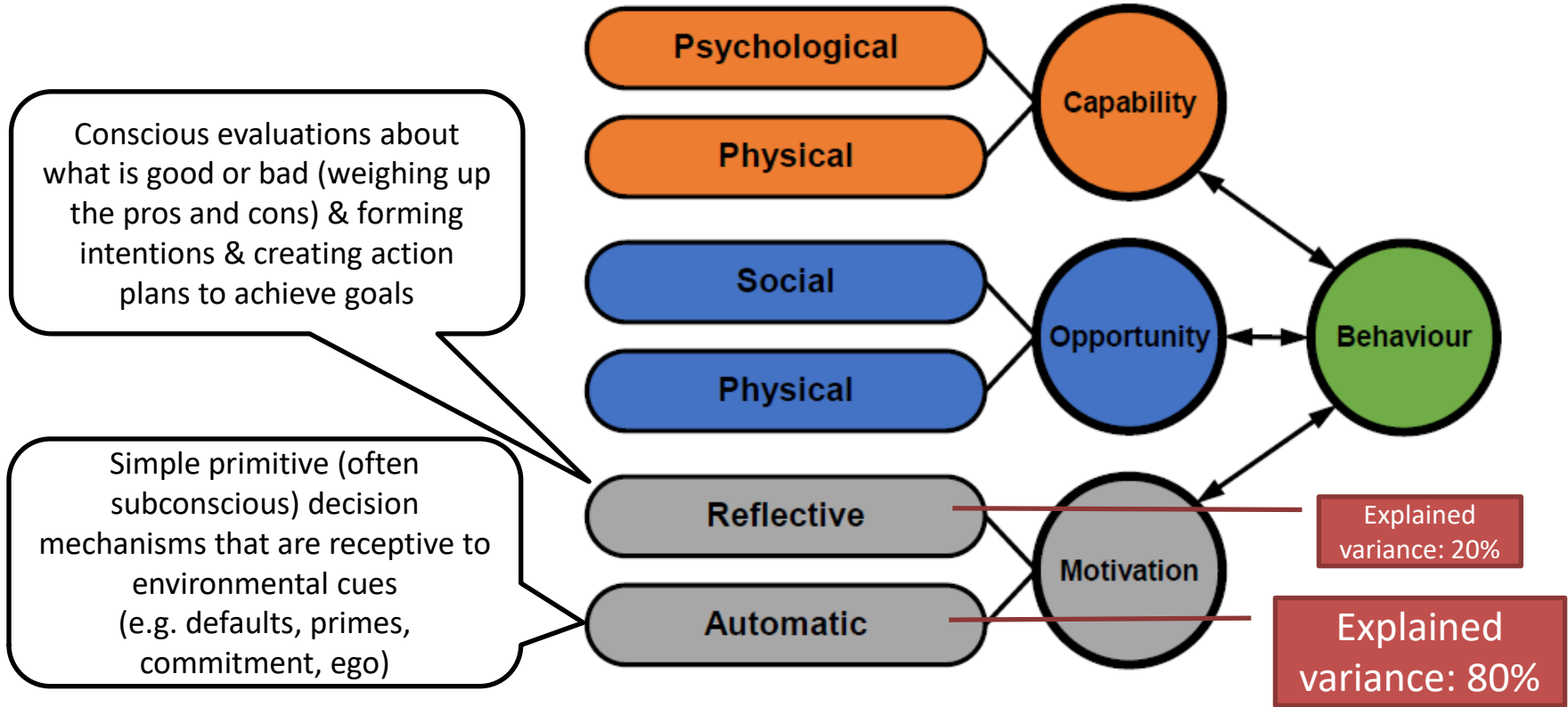


Fig. 1. Proportion of rulings in favor of the prisoners by ordinal position.

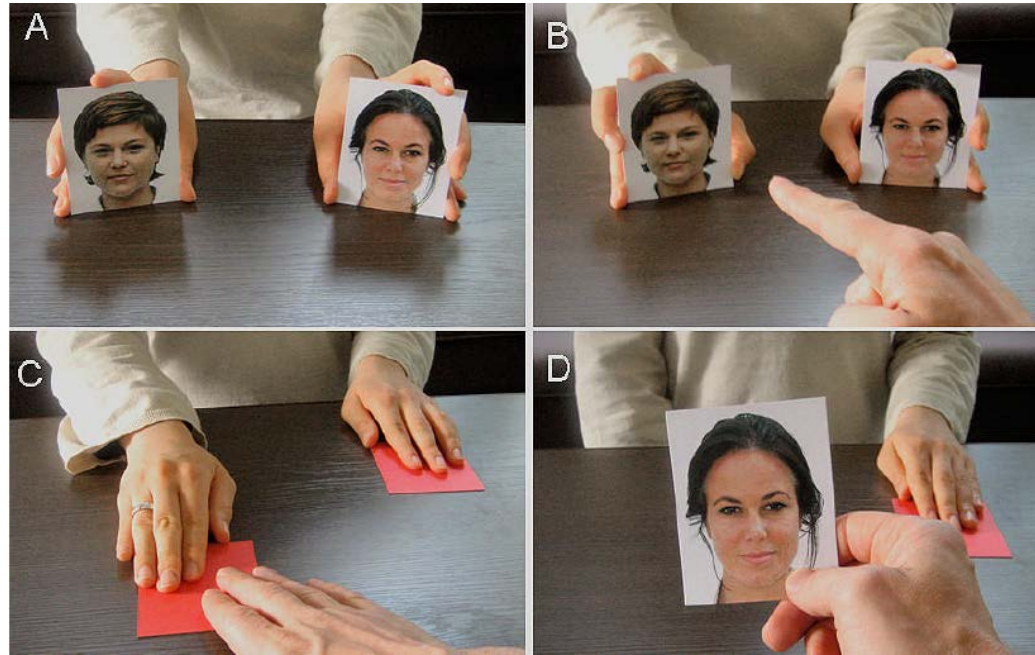




“Behavior is not so much thought about; it simply comes about”



Perceptual illusion and post-rationalisation



We gladly explain, after the fact, the choice that we did not make

Parameters analyzed:

- level of emotionality
- specificity of the explanations provided
- certainty expressed

Cognitive dissonance

This phenomenon is very well documented although the reasons for this attachment to coherence are still poorly explained

We seek for « subjective » rather than « factual » coherence

We have a strong preference for coherence...

...between our predictions and the actual events

...between our beliefs and our actions

Experimenting incoherence induces «cognitive dissonance»

Activation of automatic mechanisms that help restoring a feeling of coherence

selective forgetting

post-rationalisation

confirmation bias

Intermediate conclusion

Our judgements, our beliefs, our choices are causally influenced by...

- ...the features of situations and our rational capacity to think over it
- ...a bunch of psychological mechanisms
- ...a bunch of situational factors

} These elements are largely
beyond conscious control



The construction of our mind makes us
unaware of these causal influences
(despite available evidences)

Ingredients for nudging

One homogeneous targeted group : the “**nudgees**”



One **clearly identified change in behaviour** to be induced



A “smooth” intervention

- small change in the decision environment
- all previous choice options remain open
- easily “**resistible**”



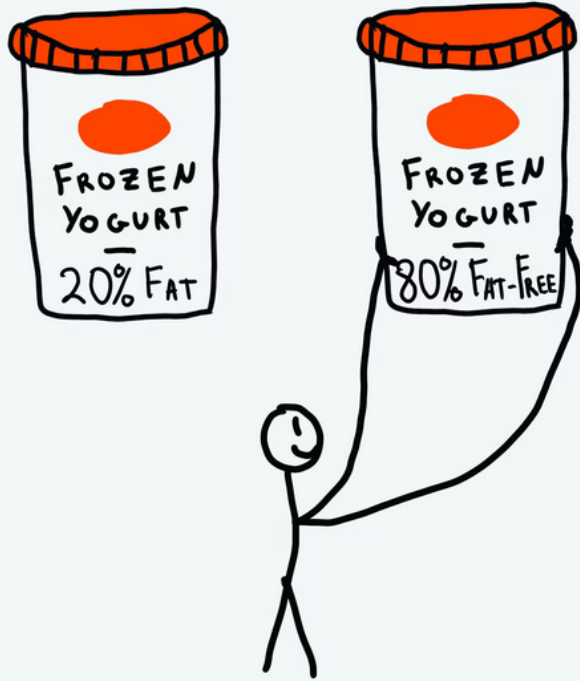
Intervention **consciously orchestrated by “nudgers”**

Good evidences that the intervention is likely to be effective

- past experiences, scientific data
- knowledge of decision-making mechanisms involved (heuristics & cognitive biases)



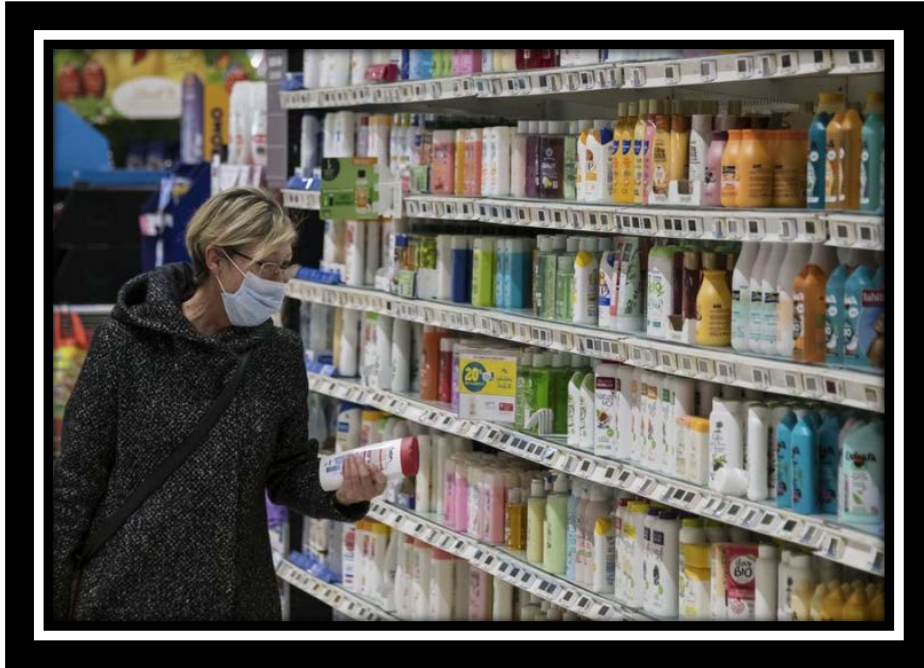
FRAMING EFFECT



© thedecisionlab.com

Our choices are influenced by the way options are framed through different wordings, reference points, and emphasis.

Decisions based on the framing effect are made by focusing on the way the information is presented instead of the information itself. Such decisions may be sub-optimal, as poor information or lesser options can be framed in a positive light. This may make them more attractive than options or information that are objectively better, but cast in a less favourable light.



Salience (eye level)



Messenger (authority – group member)

influencer





Default option



Friction

**Facilitation
(make it easy)**



There are often grey zones

Mandate



Nudging



Rational persuasion

General goal

- **promote advance care planning**

Behavioral goals

- increase the extent to which patients **think and seek information about the issue**
- **discuss it with their family and health professionals**
- **write advance directives**

Method

- use nudging techniques



Easy to understand & to make

Attractive & playful

an App to motivate & facilitate the process

Timely & relevant

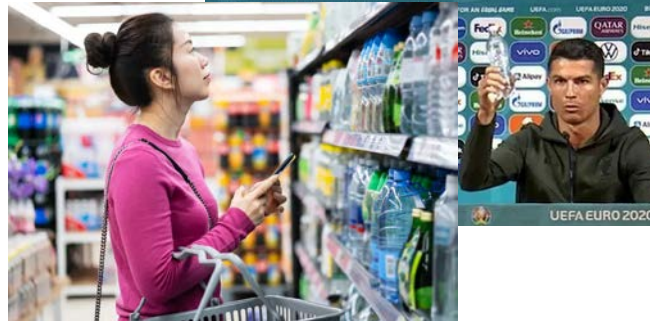
Social (reciprocity driven & commitment)



This definition of nudging is general and morally neutral

There is a clear distinction to make between

- identifying whether an intervention is a nudge or not
- evaluating whether the intervention is morally acceptable or not



Accordons-nous is a module of the app « Concerto HUG », free to download



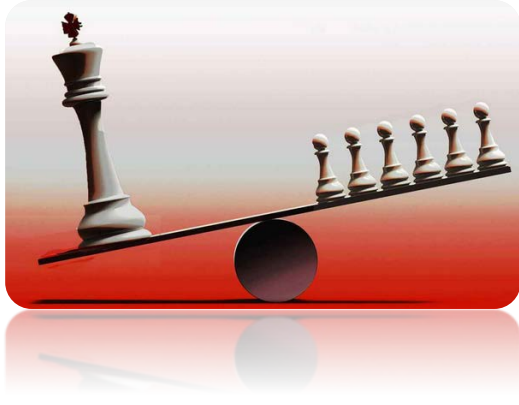
"In principle" nudging is a **soft** intervention:

- the choice options are left open (space for saying no)
- it is a benevolent

Nevertheless, nudging involves a form of paternalism: the nudgers decide the behavior they would like to promote among the population of nudges: to some extent, it is a **threat to rationality and autonomy**



Nudging involves asymmetry of power: nudgers know nudgee's psychological weakness (biases, fast decision mechanisms) and design ways to exploit these weakness in order to induce a given behavior.

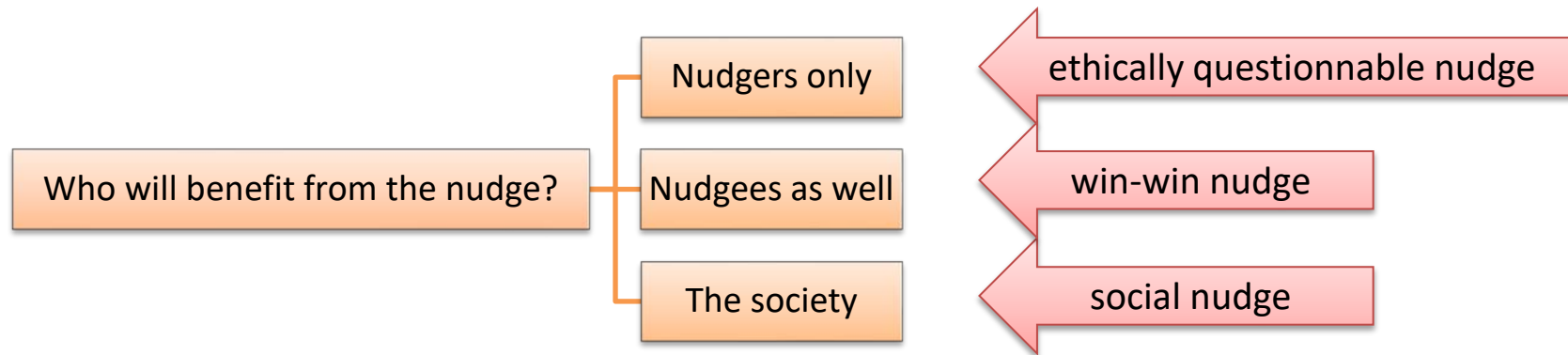


This asymmetry of power and knowledge provides **opportunities to exploit vulnerabilities.**

Nudging involves moral responsibility on part of « nudgers » (the decision-makers): the responsibility to address ethical issues such respect for autonomy and non-exploitation of vulnerability.

Before deciding to apply a nudge, it is important to make a detailed ethical evaluation.

Nudging for the benefit of whom ?



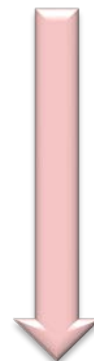


Ethics of nudges: A general framework with a focus on shared preference justifications^s

Christine Clavien

Institute Ethics History Humanities, University of Geneva, Genève, Switzerland

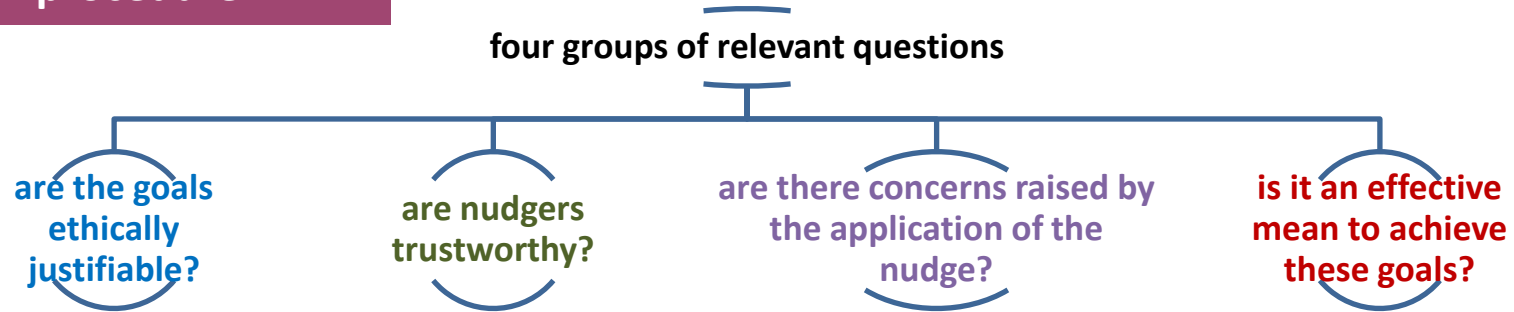
**Ethical framework
for assessing nudges**



**Nudge Evaluation Tool
for policy-makers**



A Nudge evaluation procedure



After weighing-up all the pro and contra aspects raised by this four-level investigation, one can take a balanced decision in favour or against a nudge



Four groups of relevant questions

are the goals ethically
justifiable?

are nudgers
trustworthy?

are there concerns raised by the
application of the nudge?

is it an effective mean
to achieve these goals?



Desirable
consequences



Moral
principles



Rigorous
procedure



Conflict of
interest



Side effects



Compared to status quo



Shared goals



Good
intention



Legitimacy



Capability



Autonomy



Compared to alternative

Let us use the evaluation procedure with an example

Smart glasses provided by a cantonal service to elderly citizens.

These devices include the following functions:

- sound and eye tacker
- gps localisation
- send messages

The system perceives when the aged person is stressed out or lost and can send an urgency messages to relatives with gps location.

The glasses nudge elderly person to keep their everyday practices such as go shopping, go to the gym, or walk the dog.



are the goals
ethically
justifiable?

are nudgers
trustworthy?

are there concerns raised
by the application of the
nudge?

is it an effective
mean to achieve
these goals?

What are nudgers' goals?

→ Promotion of health and autonomous living

→ Keep patients longer at home → relief for the health system

win-win nudge (but not
for the same reasons)

Ways to justify
a goal

Are the aimed consequences desirable?

Is the goal a way to fulfil important values & moral principles?

Does the goal stem from good intentions? (i.e. motivation to care for, to act in favour of)

Do stakeholders share the goal?



are the goals
ethically
justifiable?

are nudgers
trustworthy?

are there concerns raised
by the application of the
nudge?

is it an effective
mean to achieve
these goals?

Questions to
assess
trustworthiness

Are there **conflicts of interest** & are they explicitly stated and adequately addressed?

Does the decision to nudge stem from a
rigorous and transparent procedure?

systematic evaluation procedure?

collective decision including
stakeholders?

Is the nudger **capable** of implementing the nudge appropriately?

Is the nudger **legitimated** to actively promote the aimed goal ?

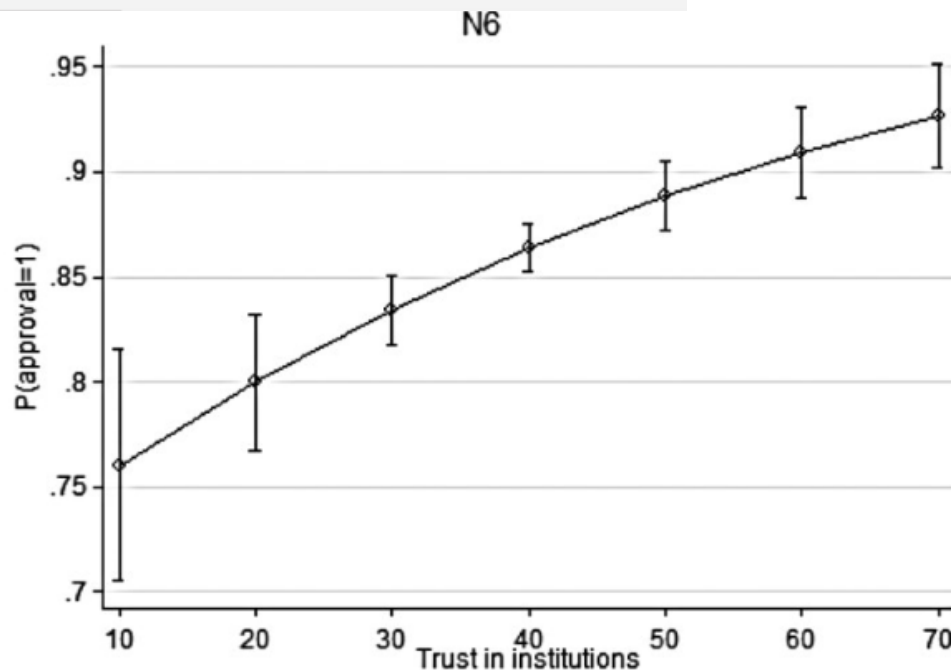


Trusting nudges? Lessons from an international survey

Cass R. Sunstein, Lucia A. Reisch  & Micha Kaiser 

Pages 1417-1443 | Published online: 11 Oct 2018

“We investigate whether people approve of a list of 15 nudges regarding health, the environment, and safety issues. A particular focus is whether trust in public institutions is a potential mediator of approval. The study confirms this correlation.”

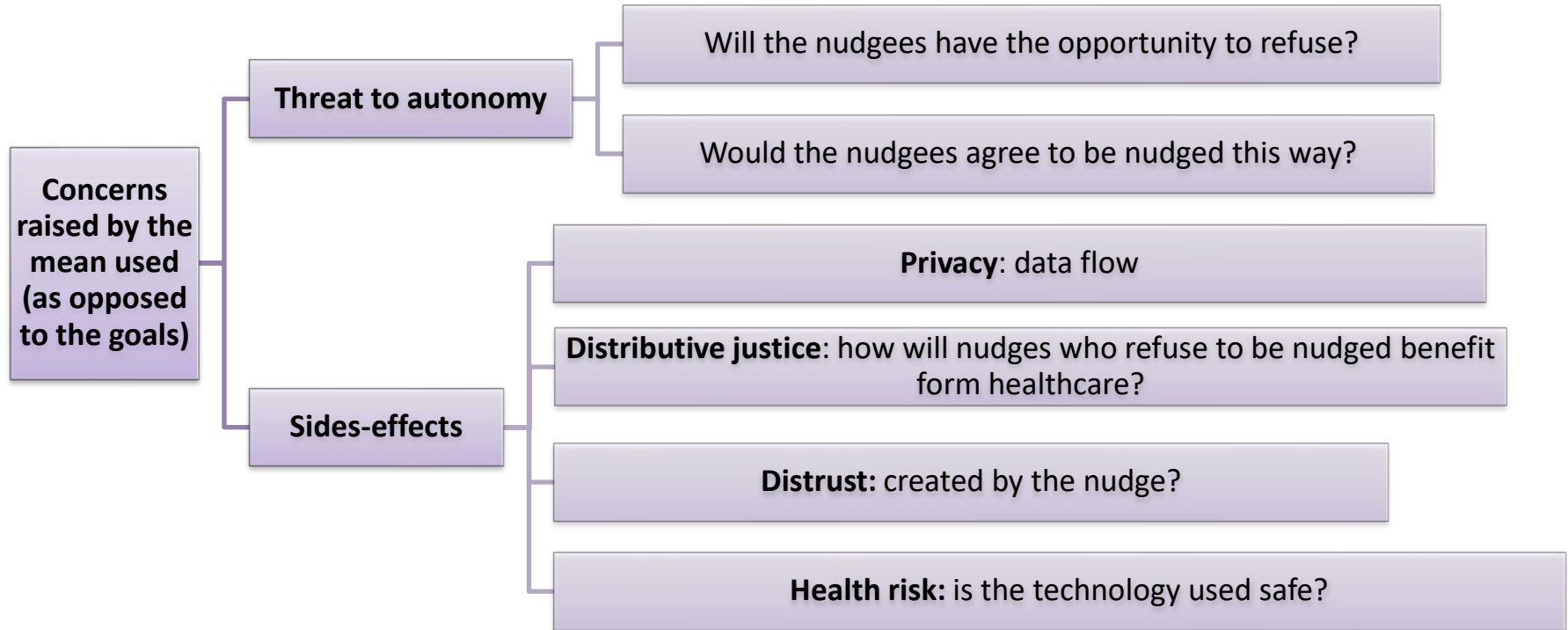


are the goals
ethically
justifiable?

are nudgers
trustworthy?

are there concerns raised
by the application of the
nudge?

is it an effective
mean to achieve
these goals?

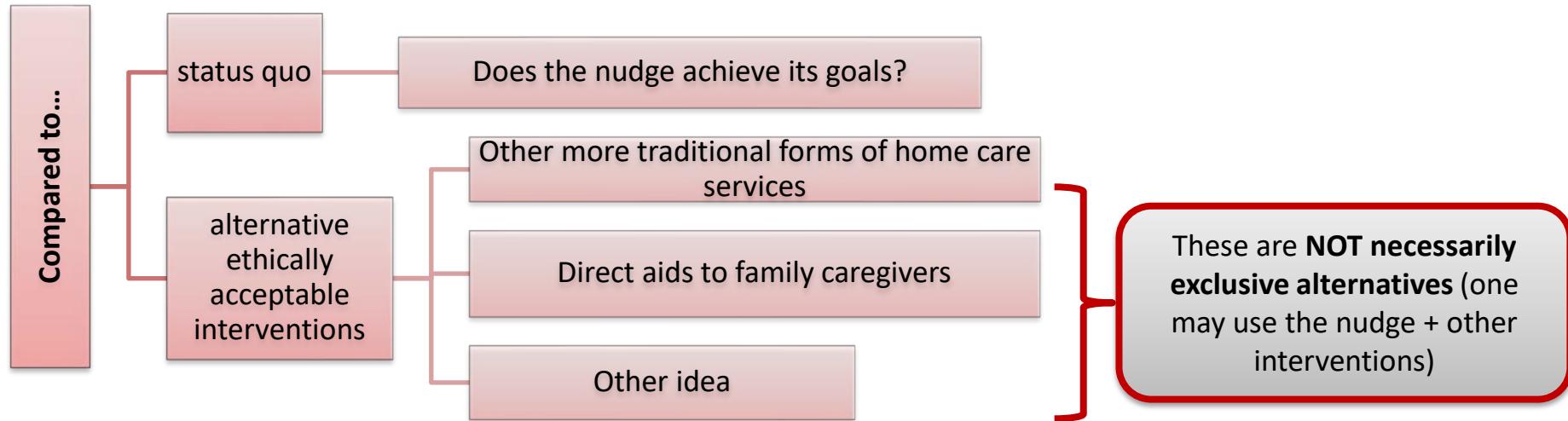


are the goals
ethically
justifiable?

are nudgers
trustworthy?

are there concerns raised
by the application of the
nudge?

is it an effective
mean to achieve
these goals?



Four groups of relevant questions

are the goals ethically
justifiable?

are nudgers
trustworthy?

are there concerns raised by the
application of the nudge?

is it an effective mean
to achieve these goals?

Desirable
consequences

Moral
principles

Rigorous
procedure

Conflict of
interest

Side effects

Compared to status quo

Shared goals

Good
intention

Legitimacy

Capability

Autonomy

Compared to alternative