# **PSYCHOLOGY AND COGNITIVE SCIENCES**

Open Journal 6



# Review

# **Nudge: Towards a Consensus View?**

Vincent Berthet, PhD1\*; Benjamin Ouvrard, PhD2

<sup>1</sup>Assistant Professor, University of Lorraine, Nancy, France

<sup>2</sup>Post Doctorate, Department of Economics, Toulouse School of Economics, Toulouse 31015, France

#### \*Corresponding author

Vincent Berthet, PhD

Assistant Professor, University of Lorraine, Nancy, France; Tel. +33 354505214; E-mail: vincent.berthet@univ-lorraine.fr

#### Article information

Received: November 16th, 2018; Revised: December 24th, 2018; Accepted: December 31st, 2018; Published: January 4th, 2019

#### Cite this article

Berthet V, Ouvrard B. Nudge: Towards a consensus view? Psychol Cogn Sci Open J. 2019; 5(1): 1-5. doi: 10.17140/PCSOJ-5-143

# ABSTRACT

This article presents a particular viewpoint on how nudge should be understood. The concept of nudge has generated considerable interest among academics and policymakers. However, ten years later, what is meant exactly by "nudge" is still a matter of debate. In fact, there is a fundamental discrepancy between the (original) narrow definition of Thaler and Sunstein (nudge in the narrow sense, NN) and the (later) broad definition of Sunstein (nudge in the broad sense, NB). These two definitions differ regarding the instrumental use of rationality failures, and accordingly whether the provision or disclosure of information counts as a nudge or not. From a pragmatic perspective, the paper argues for a position that consists of adopting the broad definition of a nudge while acknowledging several types of nudges, which we provide in an integrative view. We suggest that future research should assess the effectiveness of these different types of nudges separately.

#### Keywords

Nudge; Rationality; Information; Behavioral science; Cognitive bias; Policy intervention; Nudge effectiveness; Thaler; Sunstein.

# **INTRODUCTION**

As defined by Thaler and Sunstein¹ in their eponymous book, a nudge is a type of non-coercive intervention on behaviors/ choices. Nudging is not new (for instance, marketers have consistently nudged consumers as part of their branding, packaging and advertising strategies). Rather, the novelty of nudge lies in the theoretical body of work in which it is embedded. In fact, Thaler and Sunstein (henceforth T&S) emphasized that interest should focus on nudges serving a particular goal: promoting individual and social welfare. In this perspective, governments can use nudges paternalistically as a type of policy intervention alongside traditional policy tools (i.e., regulatory and economic measures). Moreover, as nudge interventions preserve freedom of choice; they implement the political doctrine of libertarian paternalism previously defined by Thaler and Sunstein.²

Despite the fact that nudge has been massively discussed in various academic areas, this concept remains ill-defined for at least two reasons. The first one is that most research confounds nudge as a method of behavior change (the mean of nudging) and nudge as a type of policy intervention (the goal of nudging). That confusion mainly results from T&S insisting that the focus

should be on welfare-promoting nudges (in their hands, "paternalistic nudges" is almost a pleonasm). However, as Hansen<sup>3</sup> emphasized, the notion of "nudge" is not necessarily related to that of "libertarian paternalism" so that we should adopt a technical (i.e., goal-independent) definition of a nudge. As a method of behavior change, a nudge can serve either private or public interests (paternalistic nudges). Consider the following example: if putting fruits at eye-level in a cafeteria is considered as a nudge, then putting candies at eye-level in a supermarket should also be considered as

The second reason why the concept of nudge is ill-defined is that the very definition of a nudge as a method of behavior change is still unclear and controversial. In fact, the equivocal nature of the concept, coupled with the variety of examples provided by T&S have fueled a lot a discussion about what nudges really are.<sup>3-5</sup> On one hand, finding common defining features to the large variety of nudge interventions is challenging. On the other hand, adopting a precise definition of nudge leaves eventually few examples fitting that definition.<sup>4</sup> In fact, many authors highlighted that some of the examples provided by T&S do not fit with their own definition.<sup>6</sup> Some authors have addressed that conceptual issue by amending the original definition of nudge,<sup>3,7</sup> others by distinguish-

©Copyright 2019 by Berthet V. This is an open-access article distributed under Creative Commons Attribution 4.0 International License (CC BY 4.0), which allows to copy, redistribute, remix, transform, and reproduce in any medium or format, even commercially, provided the original work is properly cited.

Review | Volume 5 | Number 1 |



ing different types of nudges.8-10

The present paper relies upon the three main definitions of nudge advanced by T&S and Sunstein. 11,12 We first distinguish between broad and narrow senses of the term "nudge". Then we argue why we should adopt a broad definition of nudge, and we finally provide an integrative view of the different types of nudges that nudge in the broad sense covers.

## NUDGE IN THE NARROW VS. BROAD SENSE

In their early work, T&S come up with two definitions of nudge:

Definition 1: "A nudge, as we will use the term, is an aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any option or significantly changing their economic incentives?".

Definition 2: "In accordance with our definition, a nudge is any factor that significantly alters the behavior of Humans although it would be ignored by Econs". 1

These two definitions (and their later refinement) describe nudging in the narrow sense (NN), revolving around the idea that nudge interventions work by harnessing rationality failures or by "exploiting cognitive biases". NN are rationally neutral (what Tor<sup>13</sup> referred to as the "rational neutrality" condition) so that definition 2 derives from definition 1.1 The rationale is as follows: under perfect information, rational agents make their choices based on their own preferences and the constraints (option set) they face. In such conditions, choices that people make reveal their preferences. As a nudge intervention leaves the rational determinants of choice (option set and economic incentives) unchanged (definition 1), it would therefore not affect the behavior of Econs (definition 2), and any change in behavior observed could only be attributed to non-rational (i.e., behavioral) factors, namely rationality failures.<sup>2</sup> Put differently, nudges use these rationality failures instrumentally<sup>14</sup>. Noteworthy, the rational neutrality is the signature of NN, which separates it from other psychological approaches to behavior change such as rational persuasion or motivation. 15 The definitions that followed that of T&S confirmed the instrumental use of rationality failures as the defining feature of nudge. 3,5,16,17

More recently, Sunstein<sup>11,12</sup> endorsed a significantly different definition of nudge:

Definition 3: "Nudges are private or public initiatives that steer people in particular directions but that also allow them to go their own way". 12

Such a large definition describes nudge in the broad sense (NB). According to Sunstein, the defining feature of nudges is not the instrumental use of rationality failures but rather the preserving of freedom of choice, in accordance with libertarian paternalism. As it means that nudges also target unbiased people (by providing them with the information they may lack), definition 3 is incompatible with definition 2. Sunstein frequently puts forward the global positioning system (GPS) device as an example of a nudge that works independently of rationality failures: a GPS steers people in a certain direction whether they are biased or not.

The difference between NN and NB is manifest regarding the issue of whether the provision or disclosure of information counts as a nudge or not. In fact, some of the examples provided by T&S consist of providing information to the individual (e.g. fuel economy labels, energy efficiency labels, calorie labels, cigarette warnings, graphic health warnings, and social norms).3 Whether the provision of information should qualify as a nudge has been controversial.3,11,18 In fact, while rational choices are determined by preferences and constraints under perfect information, rational agents are also responsive to new information under imperfect information (Econs are not necessarily perfectly informed; they just process the available information in a consistent manner). By definition, the provision of information does not forbid any option. Therefore, as long as it does not change the economic incentives either, providing information fits definition 1 while not fitting definition 2. As information provision fits NB but not NN, authors who explicitly addressed this issue hold contradictory positions.3,11 Relatedly, NB might also include what Grüne-Yanoff and Hertwig<sup>19</sup> qualify as "boosts", that is, interventions that aim to help people making better choices for themselves by overcoming their rationality failures rather than exploiting them. More broadly, such interventions are part of a "think" strategy, which in most elaborate form is deliberative democracy in which citizens engage in collective thinking and debate.20

## WHY WE SHOULD ADOPT NB

One might argue that NB appears eventually as a catchall term, which reduces the specificity and value of the concept of nudge. Notably, it could be highlighted that information provision is primarily a traditional economic policy (especially relevant when firms have an incentive to conceal information and consumers have no incentive to search for information). Information provided was actually used paternalistically way before the advent of nudges (for example, the US policy tackling obesity has mainly drawn on information provided to consumers with practices like calorie labeling). Still, we suggest that policymakers and academics should favor NB rather than NN for at least two reasons. The first reason is that nudge appears as an "uncatchable" object: even NB is insufficient to

<sup>1.</sup> For NN to be coherent, the notion of "economic incentives" in definition 1 had to be refined [3, 4, and 14]. Indeed, sticking to a meaning of economic incentives in terms of financial constraints would entail to count as nudges numerous interventions that actually alter a rational aspect of choice. Moreover, Mongin and Cozic [4] proposed that economic incentives should also include beliefs and preferences in addition to a physical set of options and financial constraints. 2. "Econs" refer to perfectly rational agents, by opposition with "Humans" who exhibit rationality failures.

<sup>3.</sup> Regarding nudges that work through information, one must distinguish between interventions that involve the disclosure of new information from interventions that consist of manipulating or translating existing information (changing the format or presentation of information without changing the content). The former are NB while the latter are NN (see below).

<sup>4.</sup> For instance, disclosure of information can generate public reputation costs thereby changing the economic incentives of firms to engage in pro-environmental behavior (e.g., the U.S. Toxic Release Inventory).



decide whether a particular intervention qualifies as a nudge or not. As outlined by Mongin and Cozic, doing this based on definition 2 ("nudge 2") is problematic as deciding whether a particular psychological feature (e.g. loss aversion, commitment, conformism) counts as a rationality failure or not can be tricky. Moreover, some interventions fit only one of the two definitions of T&S. For instance, is the tobacco package warning "Smoking causes almost 9 out of 10 lung cancers" a nudge? While it steers behaviors by harnessing a logical fallacy (thereby satisfying definition 2), in this case, the tendency to equate the conditional probabilities P(A|B) and P(B|A), such a message may change beliefs or preferences (thus violating definition 1).

The second reason for adopting NB is that when considering nudge, pragmatism should prevail. From the start, T&S put the focus on the use of nudge by governments as a public policy tool to promote social welfare. In a public policy perspective, why a message such as "most people in your local area pay their tax on time" would be considered as a nudge while "smoking causes cancer" would not be considered as one? Suppose that a government is willing to implement nudges on a particular issue, selecting only interventions that fit the definition of nudge adopted. The cost of adopting NB is to increase the rate of "false-positive nudges" (i.e., implementing interventions that actually do not harness rationality failures). The cost of adopting NN, however, seems more significant: disregarding interventions that would effectively enhance social welfare. For instance, when tackling the public health issue of smoking through channels other than regulation and economic incentives, does it really matter whether a warning such as "smoking causes cancer" harnesses a rationality failure or not? In brief, looking for a precise (unreachable?) definition of nudge might distract us from the fact that nudge is primarily about public policy, hence pragmatism and effectiveness. We should avoid the trap of abstraction and consider nudge as a broad method of behavior influence, defined by not imposing significant material costs on individuals and eventually preserving freedom of choice (definition 3).

# NUDGE TYPES: AN INTEGRATIVE VIEW

If one adopts NB, one must acknowledge a variety of nudges. We suggest that the typologies of nudges proposed so far fall into two categories, mechanism-oriented and technique-oriented. These two kinds of typologies have been advanced in order to discuss the ethics and the effectiveness of nudging, respectively. On one hand, mechanism-oriented typologies of nudges aim at identifying the psychological mechanisms nudges work by.<sup>8-10,14,21</sup> Such typologies have been introduced primarily to address the ethical issues raised by nudging.<sup>7,16,22-25</sup> Indeed, understanding the underlying psychological mechanisms of nudging is necessary to address recurrent ethical critics such as "nudges work better in the dark".<sup>16</sup> It turns out that all mechanism-oriented typologies of nudges revolve around the distinction between System 1 (automatic processes) vs. System 2 (deliberative processes).<sup>26</sup>

Baldwin<sup>8</sup> distinguished three types of nudges according to the degree to which they affect the autonomy of individuals. While first degree nudges encourage conscious deliberation, second (e.g. default rule) and third-degree nudges (e.g. framing) draw on automatic responses. The difference between second and third-degree nudges refers to transparency and is, therefore, a difference of degree rather than of kind. While embracing the broad definition of nudge, Barton, and Grüne-Yanoff' distinguished between three different types of nudges according to the nature of the interventional mechanism: heuristics-triggering and heuristics-blocking nudges on the one hand, and, on the other hand, informing nudges, which do not rely on any specific psychological shortcut. Sunstein<sup>21</sup> himself distinguished between System 1 nudges, which rely on automatic processing, from System 2 nudges, which rely on deliberative processing (while reporting that most people prefer System 2 nudges). In order to assess the effectiveness of nudges in the health domain and discuss the associated ethical issue, Lin et al10 distinguished between two types of nudges according to the amount of reevaluation of information on which people's choices are made. While Type 1 nudges are based on automatic processing (e.g. smaller plate sizes, footprints), Type 2 nudges aim to promote deliberation (e.g. calorie labeling, health warnings).

Finally, Hansen and Jespersen<sup>14</sup> also proposed a Type 1/ Type 2 nudges distinction. However, their distinction draws on what is nudged rather than how nudging is done. In their framework, Type 1 nudges influence behaviors ("non-voluntary actions") whereas Type 2 nudges influence choices (the "end-result of the intervention of reflective thinking"). Nudges can actually steer behaviors (e.g. driving speed) or bias choices (e.g. organ donation) in particular directions. However, one might wonder if distinguishing types of nudges based on what is nudged is the most relevant approach. Firstly, the frontier between behavior and choice is often fuzzy (for instance, in the examples provided by the authors, one might question why the fly-in-the-urinal nudge is classified as a Type 2 nudge while the organization of the cafeteria is classified as a Type 1 nudge). Second and most importantly, nudges that target the same behavior/choice can actually differ significantly. Let's consider the fly-in-the-urinal nudge, which prompts men to take aim. According to Hansen and Jespersen,14 this nudge qualifies as a Type 2 nudge because it involves a decision to aim for the fly or not. Now imagine that a nudger sets up a device that provides to the user real-time feedback on his accuracy at taking aim. In the framework of Hansen and Jespersen,14 such a nudge would also qualify as a Type 2 nudge since the behavior results from an intervention of reflective thinking. Still, both nudges are clearly different: one works by changing astutely the environment, the other works by providing information.

On the other hand, technique-oriented typologies of nudges were introduced to address the issue of the effectiveness of nudging that is when and why nudges work. Such typologies identify the different techniques of nudging, regardless of the underlying psychological mechanisms. The taxonomy of Münscher, Vetter, and Scheuerle<sup>27</sup> identifies nine categories of choice architecture techniques (e.g. change choice defaults, translate information, provide social reference point). Notably, Szaszi, Palinkas, Palfi, Szollosi, and Aczel<sup>28</sup> relied upon this taxonomy to estimate the effectiveness of nudging. The authors listed all empirical arti-



cles (published between January 2008 and May 2016) in which a choice architecture intervention was tested, resulting in a database containing 116 empirical articles with 156 studies and more than 422 tested interventions, which were classified according to the taxonomy of Münscher et al.<sup>27</sup> Recently, Congiu and Moscati<sup>29</sup> also proposed a technique-oriented typology by distinguishing between nudge interventions based on "Emvironment" (e.g. items arrangement, speed bumps on roads, defaults) (what might be called structural nudges) from those based on "Message" (e.g. "Smoking damages the lungs", "Most of your peers already bought it") (what might be called informational nudges). In doing so, they formulated the broadest possible technique-oriented typology of nudges (note that these two categories are not mutually exclusive: a single nudge may operate on both the Message dimension and the Environment dimension of the choice architecture).

Noting that mechanism and technique are orthogonal factors, and using the broadest dichotomy for each factor (System 1/System 2 and Environment/Message, respectively), one can provide an integrative view of nudge types (see Table 1).

Mechanism		
Technique	System I (automatic)	System 2 (deliberative)
Environment (structure)	Cafeteria re-arrangement     Chicago Lake Shore drive     Footprints	Change of printer defaults     Visual illusions in traffic     Locating a vending machine in a more remote room
Message (information)	• Warnings (e.g., "Danger!") • Cues (e.g., "Look Right" in London, "Push/Pull" on doors)	Calorie labeling     Health warnings     Social norms

Note that while such a framework might be useful to understand how nudges work and design new interventions, a priori classification can be hazardous. In fact, nudge types are more easily distinguishable in theory than in practice. Regarding the technique factor, deciding whether a particular intervention involves the provision of information or not might be tricky. On one hand, some nudges that are thought to be informational do not involve the disclosure of new information but rather the re-framing or translation of information that was already available. In fact, most nudges that involve information (e.g. reminders, feedback, information framing or translation) are actually structural nudges. On the other hand, nudges a priori categorized as structural actually do provide new information (such as defaults potentially providing information about a policymaker's attitudes.<sup>30</sup> Regarding the mechanism factor, deciding whether a particular intervention works by System 1 or System 2 can also be tricky. For instance, in some cases, providing social information is aimed at stimulating deliberation (System 2). However, providing such information may do nothing more than provide a different heuristic rule by which people can make their choice (e.g., "I'll choose the option that most people choose"). To us, the fact that categorizing nudges based on a priori thoughts can be quite difficult is an additional reason for embracing a pragmatic approach of the nudge.

# CONCLUSIONS

Since the seminal work of T&S, two phases of research can be identified in the growing field of nudging. While the first phase was primarily theory-driven by focusing on theory and philosophy, the current phase is data-driven and aims at assessing the effectiveness of nudges.<sup>31</sup> We are still unable to define what nudge exactly is and we probably won't.

Accordingly, researchers interested in nudging should stick to two heuristic guidelines. The first one is to conceptually distinguish between nudge as a method of behavior change (the mean of nudging) from nudge as a type of policy intervention (the goal of nudging). The second one is to adopt a broad definition of nudge while acknowledging several types of nudges. Those guidelines are notably suited to address the issues of the effectiveness and ethics of nudging, the ones that ultimately matter. In particular, assessing the effectiveness of the different types of nudges separately is a promising venue for further research.

#### CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

## REFERENCES |

- 1. Thaler RH, Sunstein CR. Nudge: Improving Decisions About Health, Wealth, and Happiness. New Haven, Connecticut, USA: Yale University Press. 2008.
- 2. Thaler RH, Sunstein CR. Libertarian paternalism. *The American Economic Review*. 2003; 93: 175-179. doi: 10.1257/000282803321947001
- 3. Hansen PG. The definition of nudge and libertarian paternalism: Does the hand fit the glove? *European Journal of Risk Regulation*. 2016; 7: 155-174. doi: 10.1017/S1867299X00005468
- 4. Mongin P, Cozic M. Rethinking nudge: Not one but three concepts. *Behavioural Public Policy*. 2018; 2: 107-124. doi:10.1017/bpp.2016.16
- 5. Hausman DM, Welch B. Debate: To nudge or not to nudge. *Journal of Political Philosophy*. 2010; 18: 123-136. doi: 10.1111/j.1467-9760.2009.00351.x
- 6. Bonell C, McKee M, Fletcher A, Wilkinson P, Haines A. One nudge forward, two steps back. *BMJ*. 2011; 342: d401. doi: 10.1136/bmj.d401
- 7. Saghai Y. Salvaging the concept of nudge. *Journal of Medical Ethics*. 2013; 39: 487-493. doi: 10.1136/medethics-2012-100727
- 8. Baldwin R. From regulation to behaviour change: Giving nudge the third-degree. *Modern Law Review.* 2014; 77: 831-857. doi: 10.1111/1468-2230.12094



- 9. Barton A, Grüne-Yanoff T. From libertarian paternalism to nudging—and beyond. *Review of Philosophy and Psychology.* 2015; 6: 341-359. doi: 10.1007/s1316
- 10. Lin Y, Osman M, Ashcroft R. Nudge: Concept, effectiveness, andethics. *Basic and Applied Psychology*. 2017; 39: 293-306. doi: 10.1080/01973533.2017.1356304
- 11. Sunstein CR. Nudges, agency, navigability, and abstraction: A reply to critics. Review of Philosophy and Psychology. 2015; 6: 511-529.
- 12. Sunstein CR. Misconceptions about nudges. *Journal of Behavio-ral Economics for Policy*. 2018; 2: 61-67. doi: 10.2139/ssrn.3033101
- 13. Tor A. The critical and problematic role of bounded rationality in nudging. In: Mathis K, Tor A (Eds). *Nudging. Possibilities, Limitations and Applications in European Law and Economics.* Switzerland: Springer International Publishing. 2016.
- 14. Hansen PG, Jespersen AM. Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behavior change in public policy. *European Journal of Risk Regulation*. 2013; 1: 3-28.
- 15. Miller DT, Prentice D. Psychological levers of behavior change. In: Shafir E (Ed). *The Behavioral Foundations of Policy*. New Jersey, USA: Princeton University Press. 2013.
- 16. Bovens L. The ethics of nudge. In: Grüne-Yanoff T, Hansson SO (Eds). *Preference Change: Approaches from Philosophy, Economics and Psychology*. Berlin, Germany: Springer. 2009.
- 17. Oliver A. From nudging to budging: Using behavioural economics to inform public sector policy. *Journal of Social Policy*. 2013; 42: 685-700. doi: 10.1017/S0047279413000299
- 18. Schubert C. Green nudges: Do they work? Are they ethical? *Ecological Economics*. 2017; 132: 329-342. doi: 10.2139/ssrn.2729899
- 19. Grüne-Yanoff T, Hertwig R. Nudge *versus* boost: How coherent are policy and theory? *Minds and Machines*. 2016; 26: 149-183. doi: 10.1007/s1102
- 20. John P, Smith G, Stoker G. Nudge nudge, think think: Two strategies for changing civic behaviour. *Political Quarterly*. 2009; 80:

- 361-370. doi: 10.1111/j.1467-923X.2009.02001.x
- 21. Sunstein CR. People prefer system 2 nudges (kind of). *Duke Law Journal*. 2016; 66: 121-168. doi: 10.2139/ssrn.2731868
- 22. Blumenthal-Barby JS, Burroughs H. Seeking better health careoutcomes: The ethics of using the "nudge". *Am J Bioeth.* 2012; 12(2): 1-10. doi: 10.1080/15265161.2011.634481
- 23. Ashcroft RE. Doing good by stealth: Comments on 'salvaging the conceptof nudge'. *J Med Ethics*. 2013; 39(8): 494. doi: 10.1136/medethics-2012-101109
- 24. Nys TRV, Engelen B. Judging nudging: Answering the manipulation objection. *Political Studies*. 2017; 65,199-214. doi: 10.1177%2F0032321716629487
- 25. Smith NC, Goldstein DG, Johnson EJ. Choice without awareness: Ethical and policy implications of defaults. *Journal of Public Policy & Marketing*. 2013; 32: 159-172. doi: 10.1509/jppm.10.114
- 26. Stanovich KE, West RF. Individual difference in reasoning: implications for the rationality debate? *Behavioral and Brain Sciences*. 2000; 23: 645-726. doi: 10.1017/S0140525X00003435
- 27. Münscher R, Vetter M, Scheuerle T. A review and taxonomy of choice architecture techniques. *Journal of Behavioral Decision Making*. 2015; 29: 511-524. doi: 10.1002/bdm.1897
- 28. Szaszi B, Palinkas A, Palfi B, Szollosi A, Aczel B. A systematic scoping review of the choice architecture movement: Towards understanding when and why nudges work. *Journal of Behavioral Decision Making*. 2017; 31: 355-366. doi: 10.1002/bdm.2035
- 29. Congiu L, Moscati I. Message and environment: A framework for nudges and choice architecture. *Behavioural Public Policy*. 2018; 1-17. doi: 10.1017/bpp.2018.29
- 30. McKenzie CR, Liersch MJ, Finkelstein SR. Recommendation-simplicit in policy defaults. *Psychol Sci.* 2006; 17(5): 414-420. doi: 10.1111/j.1467-9280.2006.01721.x
- 31. Benartzi S, Beshears J, Milkman KL, et al. Should governments invest more in nudging? *Psychol Sci.* 2017; 28(8): 1041-1055. doi: 10.1177/0956797617702501