



**HAL**  
open science

## Foundations for an Ontology of Nudges

Adrien Barton

► **To cite this version:**

Adrien Barton. Foundations for an Ontology of Nudges. 12th International Conference on Biomedical Ontologies (ICBO 2021), Sep 2021, Bozen-Bolzano, Italy. hal-03512325

**HAL Id: hal-03512325**

**<https://hal.science/hal-03512325>**

Submitted on 5 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Foundations for an Ontology of Nudges

Adrien Barton <sup>1,2</sup>

<sup>1</sup> *Institut de recherche en informatique de Toulouse (IRIT), CNRS, France*

<sup>2</sup> *Groupe de recherche interdisciplinaire en informatique de la santé (GRIIS), Université de Sherbrooke, Québec, Canada*

## Abstract

Introduced by Cass Sunstein and Richard Thaler, nudges have become an important tool in public policy, in particular for health-related interventions. This paper provides foundations to represent nudges in OBO Foundry ontologies. It defines a nudge as a function of an environment (or part thereof) that is designed to influence behavior, in order to contribute to the influencee's welfare or the society's welfare, and that is easily resistible. Nudges are distinguished from nudge-setting processes and nudge-activation processes.

## Keywords

Nudge, Disposition, Function, Intention, Behavior, OBO Foundry

## 1. Introduction

The domain of health sciences extends beyond biomedical sciences: an important class of health interventions consists in organizing the environment in order to influence people to behave in ways that are beneficial for their health, based on insights in behavioral sciences. Such interventions encompass in particular nudges, which have been defined by Thaler & Sunstein ([1], p. 6) as:

DEF<sub>TS1</sub>: = “[...] any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives.”

Health-beneficial nudges encompass, for example: providing smaller plates, thus decreasing the amount of food consumed (named “PLATE” thereafter [2]); rearranging cafeterias to lead people to the salad buffet before the burger stand (“CAFETERIA” [1]); or putting in place scary or disgusting messages on tobacco packs to influence people to smoke less (“TOBACCO” [3]). Nudges are not restricted to the health domain, and can promote, e.g., financial welfare of the nudgees [4] or ecological benefits [5].

The OBO Foundry [6] is to date one of the most significant attempts to build interoperable ontologies in the biomedical domain. It is mostly founded on the upper-level realist ontology Basic Formal Ontology (BFO) [7], which aims at formalizing the most general classes on which domain ontologies should be based. Recently, OBO Foundry-based ontologies have been increasingly interested in incorporating entities analyzed by behavioral sciences. In particular, the Behaviour Change Intervention Ontology (BCIO) [8] is being developed. It would thus be useful for such ontologies to encompass categories that can account for nudges and related entities. This requires to clarify the ontological nature of nudges in terms of general ontological categories proposed by BFO and OBO Foundry ontologies, and propose a solid and systematic definition of nudges. This investigation should also be of more general interest for nudge experts outside the ontological community, as it contributes to the goal of proposing a unified definition and characterization of nudges, from an applied ontology perspective which seems to be currently absent from the domain of nudge studies.

---

International Conference on Biomedical Ontologies 2021, September 16–18, 2021, Bozen-Bolzano, Italy

EMAIL: [adrien.barton@irit.fr](mailto:adrien.barton@irit.fr)

ORCID: 0000-0001-5500-6539



© 2021 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

This paper will follow the following structure. The next section presents the OBO Foundry-compliant methodology followed in this paper, as well as former works on the topic of definition and classification of nudges. A third section proposes a classification of nudges and related entities inside the BFO taxonomy, and a definition of nudges on this basis. A fourth section discusses how nudges can be further subclassified and differentiated. A conclusion follows.

## 2. Methods and Former Works

This section presents the methodology that will be followed as well as former works on the definition and classification of nudges.

### 2.1. Methods

Two methodological aspects will be of central importance: aligning new classes within pre-existing ontological categorizations and providing Aristotelian definitions.

#### 2.1.1. Ontological categorization

A goal of this paper is to classify nudge-related classes in pre-existing taxonomies, more specifically those of OBO Foundry ontologies, which include the upper ontology BFO (Basic Formal Ontology). BFO [7] introduces the two high-level categories of *Occurrent* and *Continuant* (largely shared with other upper ontologies such as DOLCE, GFO or UFO, under different names). Occurrents are entities that extend in time by having temporal parts, whereas continuants are entities that persist through time by being wholly present at each time at which they exist. The latter are divided between independent continuants, namely entities whose existence does not depend on the existence of another entity (such as material objects); and dependent continuants, whose existence depends on the existence of another entity (such as properties). Among the latter, specifically dependent continuants (e.g. the shape of an object) cannot migrate from one bearer to another, whereas this is possible for generically dependent continuants (e.g. information content entities). Specifically dependent continuants encompass qualities and realizable entities, where the latter can be realized in specific processes. Realizable entities include dispositions, which can be intuitively identified with causal factors ([7], [9], [10]). Finally, an important subclass of *Disposition* is *Function*, which encompasses dispositions that exist in virtue of a physical make-up of the bearer that “came into being either through evolution (in the case of natural biological entities) or through intentional design (in the case of artifacts), in order to realize processes of a certain sort.” [11]

On top of inserting nudge-related entities in such taxonomies, this paper will aim at formulating a few formal axioms satisfied by the relevant classes. It is not always possible to do so, as it would sometimes require constructs that are not yet introduced in BFO and OBO Foundry ontologies; in such cases, we will refrain from introducing axioms (which is compatible with the open-world approach of OWL ontologies) and simply describe such characteristics in natural language definitions.

#### 2.1.2. Aristotelian Definitions

A good practice, in writing ontological definitions, is to use so-called “Aristotelian definitions” ([6], [12]), on the model of Aristotle’s definition of “human” as an animal that is rational. Such definitions take the “genus-species” form: an A is defined as a B that C’s, where B is a parent class of A, and C is the difference characterizing the instances of B which are also instances of A. Thus, we will aim at proposing Aristotelian definitions of key central terms, in particular “nudge”.

An important goal of this paper will be to propose a definition of nudges that can be used in ontologies. A definition of nudges should arguably strive at: *being precise* – that is, to use well-defined terms; *using a consensual vocabulary* – namely, to use terms whose definition can be accepted by a large part of the scientific community, rather than heavily debated constructs; and *being exhaustive* –

that is, that it can apply to all recognized instances of nudges. Finally, a definition of nudges should also strive, as much as possible, at being *specific* – that is, to exclude from its range constructs that are classically not recognized as nudges by the scientific community. However, this requirement is arguably less important than the others: it might be the case, when devising a definition of nudges satisfying the three first requirements, that it would lead to a definition encompassing interventions that are not always identified as nudges by the whole scientific community, but only some members of it.

## 2.2. Former works

Two strains of academic works are especially relevant for the current project: works that aim at defining nudges (and contrast them with germane categories); and works that aim at classifying the different subkinds of nudges that have been proposed or implemented (note that those two strains are only a subset of the philosophical or conceptual analysis literature on nudges, as the bulk of it is rather interested in the ethical and political justifications of nudges).

Concerning the first strain, we will contrast a few major definitions that have been proposed in the literature, without engaging into a systematic review. On top of definition  $DEF_{TS1}$ , several definitions have been proposed in the literature (see [13], [14] for overviews):

- $DEF_{TS2}$ : “a nudge is any factor that significantly alters the behavior of Humans although it would be ignored by Econs” ([1], p. 8)
- $DEF_S$ : “Private or public initiatives that steer people in particular directions but that also allow them to go their own way.” ([15], p. 61)
- $DEF_H$ : “A nudge is ... any attempt at influencing people’s judgment, choice or behavior in a predictable way which works by making use of [people’s] boundaries, biases, routines and habits as integral parts of such attempts” [16]
- $DEF_{MC2}$ : “an intervention that uses rationality failures instrumentally.” (named “nudge 2” in [17])
- $DEF_{MC3}$ : “a welfare-promoting intervention that tries to reduce the negative effects of rationality failures.” (named “nudge 3” in [17])
- $DEF_{CM}$ : “nudges are interventions that attempt to influence people’s behavior by exploiting, at least in part, their rationality failures, and that do not increase, exclusively, the well-being of nudgers” [14]

Although each of those definitions illuminate some aspects of the nature of nudges, they arguably do not satisfy all criteria mentioned above in section 2.1.2.  $DEF_{TS1}$  and  $DEF_{TS2}$  do not account for the intentional aspect of nudges – an important aspect, as argued below in 3.3.  $DEF_{TS2}$  is not systematic: as explained by Congiu & Moscati [14], it excludes very minor change of prices whose effects are mostly psychological rather than a change of economic incentives – and that could not, strictly speaking, be fully ignored by Econs. The entity “Econ” is also arguably an abstraction difficult to define (see [17] for more criticisms of this definition).  $DEF_S$  may lack precision (what does it mean exactly to “allow people to go their own way”?) and specificity (some changes in economic incentives that would commonly not count as nudges arguably still allow “people to go their own way”).  $DEF_H$  may lack exhaustivity (why is the term “heuristic” not included in the list “boundaries, biases, routines and habits”?). The vocabulary used by  $DEF_{MC2}$ ,  $DEF_{MC3}$  and  $DEF_{CM}$  might not be fully consensual: they define nudges on the basis of rationality, but what is or is not rational is the object of intense debate in the scientific community (cf. [18] for an overview).

The first barrier to the integration of such definitions in an ontology, however, is that those definitions do not aim at clearly identifying what kind of general entity a nudge is, in a general metaphysical framework (see discussions in [17] and [19] though, to which we will come back in section 3.3): nudges are characterized as “factors”, “interventions”, “initiatives” or “attempts”, but as we will see, such categories may be ambiguous. Our starting point will thus be to propose a classification of nudges that fits within BFO’s framework.

To conclude this section, note that a second strain of works is related to our general investigation, namely works that classify the different subkinds of nudges (see e.g. the classification of [20]) – we will come back to this point in section 4.

### 3. An Ontological Analysis of Nudges

Figure 1 below presents a taxonomy of the relevant entities that will be introduced and presented in this paper.

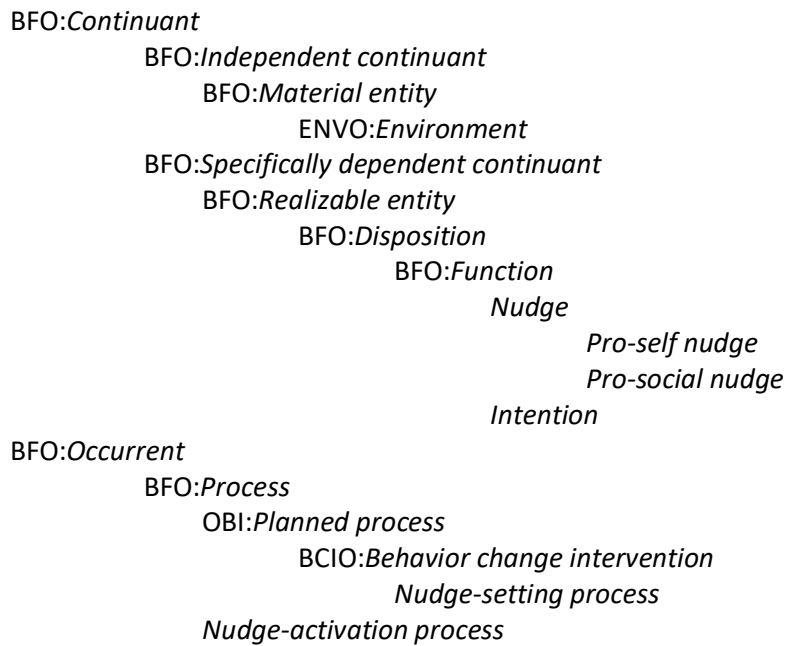


Figure 1: Taxonomy

#### 3.1. Nudge as a Disposition

Let's first consider the most general question: is a nudge an occurrent or a continuant? Contrarily to a continuant, an occurrent only exists partially at every time it exists. But a nudge arguably exists fully at every time it exists (consider the examples PLATE, CANTEEN and TOBACCO). Therefore, a nudge is a continuant. In particular, a nudge should not be confused with the process of arranging the environment in a given way that creates this nudge, which we call a "nudge-setting process": once the arrangement is established, the nudge-setting process does not exist anymore, but the nudge still does. It should also not be confused with the process of this nudge influencing the decisions of a nudgee, which we might call a "nudge-activation process": when no nudgee is currently influenced, there is no nudge-activation process, but the nudge can still exist.

The next question is whether a nudge is an independent continuant or a dependent continuant. Nudges are for sure related with independent continuants (think about PLATE); however, they cannot be identified with such independent continuants. In PLATE, the nudge is not the plate itself, but a property of the plate – related to its size. In CANTEEN, the nudge cannot be identified with the burgers, the salads or the furniture that support them; rather, the nudge is related to the *arrangement* of the food display within the cafeteria supported by those material objects. Since this arrangement depends ontologically on those material objects, it is a dependent continuant. This is in line with Thaler & Sunstein's original definition of a nudge being an "aspect" of the choice architecture. Note also that the very same nudge cannot migrate from a bearer to another. Thus, a nudge is a specifically dependent continuant rather than a generically dependent continuant.

Let us now investigate which kind of specifically dependent continuant a nudge would be. What matters in a nudge is how it affects people’s behavior causally (a nudge “alters”, “affects”, “steers” behavior). Moreover, there is a dispositional dimension in such a causal power. Indeed, suppose that a choice architect puts into place a nudge in an environment, but that nobody is currently being influenced by it: it would still exist nonetheless (we can also imagine the more extreme case in which nobody would ever interact with this nudge). Thus, there is a dispositional dimension of nudges: a nudge pushes people to behave in a certain way (a process that can be identified with the realization of the disposition) in some circumstances. Therefore, we will classify *Nudge* as a subclass of *Disposition*. In particular, a nudge-activation process can be defined as a realization (in the sense analyzed in [7], [9]) of a nudge:

*Nudge-activation process* EquivalentTo (BFO:realizes some *Nudge*)

Note that dispositions are closely linked with qualities, as most (if not all) dispositions have some qualities forming their “categorical bases” [10]. For example, in PLATE, the nudge exists because of the size of the plate, which is its categorical basis. In CANTEEN, the nudge exists because of the spatial arrangement of salads and burgers in the cafeteria, which forms its categorical basis.

### 3.2. The Bearer of a Nudge

In BFO’s framework, every specifically dependent continuant (including dispositions) has a bearer. We can assume very generally that the bearer of a nudge is an environment or a part of an environment. The OBO Foundry ENVO (Environment Ontology) [21] defines an environment as “A system which has the disposition to environ one or more material entities”, but specifies that this class is under development, so we leave here their precise ontological nature open. Thus, we will merely speak of nudges as being “environmental dispositions”<sup>2</sup>.

Note that if we follow Thaler & Sunstein in considering a nudge as an aspect of a choice architecture, the bearer of a nudge might be a choice architecture. The notion of “choice architecture” is ultimately related to a role – namely, the role of an environment to be such that agents are expected to make choices in it. Therefore, a class *Choice architecture* might be introduced as a defined class of being an *Environment* (or a part thereof) that has a *Choice architecture role*. We leave such investigation for future works, when the notion of environment will be ontologically clarified.

### 3.3. Intentional Design: Nudge as a Function

An important dimension of nudges is their intentional component. In rare cases, nudges have been defined as non-intentional entities, and merely identified with choice architectures [23]. If we would follow this conception, an environment that was not intentionally designed to influence people’s choice, but that nonetheless *does* influence people’s choice, would be considered as bearing a nudge. However, nudges have more commonly been considered as encompassing an intentional component (e.g. [24], [25]). In this spirit, they have sometimes been defined as an “intervention on the choice architecture that would alter people’s behavior” [19]. Thus, they need to be differentiated from dispositions of the environment that influence behavior without being the result of an intentional action.

What is an *intervention* on a choice architecture? We can identify at least two meanings of this term: a process, or the result of such a process. Consider the CANTEEN nudge. The intervention process here would be setting up the salad bar at the entrance and the hamburger stand at the back. But as explained earlier, this process (which is an occurrent) is not a nudge, but rather what we called a “nudge-setting process”; and the nudge would be an output of this process.

We can differentiate nudges from non-intentional behavior-influencing environmental dispositions as follows. A nudge appears as the result of some intentional actions made by a choice architect. We endorse Toyoshima et al.’s view [26] that intentions are dispositions that can be realized by some action; thus, we introduce the subclass *Intention* of BFO:*Disposition*. In the case of nudges, those actions are

---

<sup>2</sup> Note however that we adopt here a broad definition of “environment”, that encompasses not only what Congiu & Moscati [22] call “environment”, but has also what they call the “message” – such as tobacco health warnings in the nudge TOBACCO.

planned processes that result in the existence of the nudge. That is, using classes from OBO Foundry's ontologies COB (Core Ontology for Biology and Biomedicine) and OBI (Ontology for Biomedical Investigations) [27], we could characterize a nudge as the result (the specified output) of some intentional action, which is what we called before a "nudge-setting process":

*Nudge* SubClassOf (COB:specified\_output\_of some *Nudge-setting process*)  
*Nudge-setting process* SubClassOf (realizes some *Intention*)

More precisely, a nudge is the output of a proximal intentional action of setting up a nudge, where the distal intention behind this action is to influence behavior<sup>3</sup>. Therefore, *Nudge-setting process* is classified as a subclass of BCIO:*Behavior change intervention*, which is defined as "An intervention that has the aim of influencing human behavior" [8]. And because of this intentional character of nudges, we can classify *Nudge* as a subclass of BFO:*Function* (which is itself a subclass of BFO:*Disposition*). Indeed, a nudge came into being "through intentional design" (in a nudge-setting process), "in order to realize processes of a certain sort" (namely nudge-activation processes), and thus satisfies BFO's definition of function.

Note that Cozic & Mongin [17] discuss whether nudges should be interpreted as "causal factors" or "interventions", where they define the latter as processes ("a nudge intervention consists in using a nudge factor to exert an influence on the choices" – a definition that make them akin to what we called "nudge-activation process"). They opt for the second interpretation, arguing that "it is not the same for an individual to be influenced by a factor in the choice conditions when someone uses this factor for a purpose and when the factor is just there without anyone in particular being in control". Barton & Grüne-Yanoff [19] make a similar choice, although they instead interpret "intervention" as what we call here a "nudge-setting processes". However, we have here shown that we can opt for the first interpretation of nudges as causal factors and still account for the intentionality of nudges, by defining nudges as functions *that are the results of intentional interventions* (namely nudge-setting processes).

We can now examine what particularizes nudges among intentional behavior-influencing functions.

### 3.4. Nudges' Easy Resistibility

T&S's definition mentions that a nudge should not "[forbid] any options or significantly [change] their economic incentives." More generally, a central characteristic of nudges that has been commonly endorsed is their easy resistibility (some authors prefer to speak of "easy avoidability" [1] or "easy reversibility" [30]). This is what distinguishes a nudge from a "shove" [31], the latter being not easily resistible.

This condition of easy resistibility has been analyzed by Saghai as involving "substantial noncontrol" [33], which he defines as follows: "A's influence to get B to  $\phi$  is substantially noncontrolling if B is able to effortlessly oppose the pressure to get her to  $\phi$  if she does not want to  $\phi$ ." This implies that the agent should have some special cognitive capacities, which have been analyzed by Saghai [33] as attention-bringing (being able to notice the nudge) and inhibitory capacities (being able to refrain to follow the nudge influence). This also implies, from an ontological point of view, that a nudge is not a sure-fire disposition, but a tendency (see [34]–[36] for some thoughts on tendencies). Following this definition, an easy resistible intervention does not forbid any option nor significantly change the economic incentives; thus, clauses excluding those two latter aspects do not need to be added to the definition as long as easy resistibility is mentioned.

Note that some features of nudges, such as their transparency (see Bovens' analysis [32]), may help to bring nudges to the attention of the nudge. However, this is not a defining feature of nudges – but rather a feature that mitigates some ethical problems typically caused by nudges.

---

<sup>3</sup> Some authors have introduced a distinction between influencing behavior and influencing choice ([28], [29]), but we refrain from making such a distinction here. As argued in [13], "the frontier between behavior and choice is often fuzzy". Moreover, all nudges ultimately aim at influencing behavior.

### 3.5. The Goal of Nudges

Not all easily resistible intentional behavior-influencing environmental dispositions are nudges. Indeed, nudges must be distinguished from germane interventions such as marketing techniques. As a matter of fact, the novelty of Sunstein & Thaler’s approach does not consist in proposing a new kind of behavior-influencing tool (as such tools have been used in marketing for some time already), but in extending the use such techniques to new domains, typically (or maybe exclusively, as we will argue) those contributing to the welfare of the influencees or of the society.

Congiu & Moscati [14] propose that an intervention counts as a nudge if it “does not increase the nudger’s well-being *exclusively*.”. But this definition might be too general, as it encompasses interventions that are devised to contribute uniquely to the welfare of a third-party who is unrelated to the nudgee (consider e.g. an advertisement that is designed pro-bono by a nudger in favor of a third-party) – and which, arguably, should not count as nudges.

Thus, we add instead the constraint that a nudge should aim (at least in part) at increasing the nudgee’s welfare or the society’s welfare (we leave open here the definition of the term “society”), which is in line with the classification of nudges as being either pro-self or pro-social by Hagman et al. [37] (see section 4 below).

On this basis, we can propose the following Aristotelian definition for *Nudge*:

*Nudge*: A function of an environment (or part thereof) that is designed to influence behavior, in order to contribute to the influencee’s welfare or the society’s welfare, and that is easily resistible.

## 4. Discussion: Subcategorizing Nudges

Once nudges are aligned with OBO Foundry ontologies, the next step will be to subcategorize them. Several proposals have considered this matter. This section will provide only a short overview of this question (see [13] for a more detailed synthesis). At least two kinds of classification systems can be used to categorize nudges: goal-oriented classifications, and mechanism-oriented classifications.

As specified above, nudges can aim at either contributing to the welfare of the nudgee, or benefitting to the society. This has been distinguished by Hagman et al. [37] under the name “pro-self” and “pro-social” nudges. We can define such classes in the current context as follows (note that those two classes are not exclusive from each other, as a nudge can contribute to both the welfare of the nudge and the welfare of society):

- *Pro-self nudge*: A nudge that aims at increasing the welfare of the nudgee.
- *Pro-social nudge*: A nudge that aims at increasing the welfare of society.

Mechanisms under which nudges operate are crucially important – in particular for their ethical analysis [19]. Note that the definition of nudges proposed here is very large. It includes what Barton and Grüne-Yanoff [19] call “heuristics-triggering” nudges (as illustrated e.g. by PLATE, CANTEEN or TOBACCO) and “heuristics-blocking” nudges (such as mandatory cooling-off periods before an important expanse, or TOBACCO health warnings that aim at blocking addictive behavior). It also include so-called “informing nudges” (following the characterization proposed by Sunstein, 2015, who uses the analogy with a GPS). The latter category might not be what most people have in mind when they discuss nudges. However, it is relevant to encompass it, as the border between informing and heuristics-triggering is sometimes fuzzy (consider a message such as “67% of XYZ University students have 4 or fewer drinks when they party” [38], which can inform but also trigger a social conformity heuristics)<sup>4</sup>.

More generally, most mechanism-focused classifications revolve around the distinction between System 1 (automatic processes) vs. System 2 (deliberative processes) [13]. Baldwin [40] distinguishes three types of nudges depending on the degree to which they affect the autonomy of individuals: first-degree nudges that foster System 2, and second- and third-degree nudges that draw on System 1 (with

---

<sup>4</sup> Nudges that belong to several categories might thus be analyzed as complex dispositions composed by several simpler dispositions (using e.g. the tools presented in [39]) which belong each to a different category.



a different between second and third-degree nudges that relies on their transparency). Sunstein [41] distinguished between System 1 nudges from System 2 nudges, where the latter rely on deliberative processing. In the same spirit, Lin et al. [42] distinguished between type 1 and type 2 nudges. Hansen and Jespersen [29] also use the vocabulary “type 1” vs. “type 2” nudges, but with different definitions: the former influence behaviors (“non-voluntary actions”), when the latter influence choices (the “end-result of the intervention of reflective thinking”) – see [13] though for a criticism of this classification. Finally, Münscher et al. [20] provide a more detailed classification of nine nudge subcategories (in three groups). Future works should identify which of those classifications are the most promising.

## 5. Conclusion

This paper has proposed a classification of nudges-related entities among BFO and OBO Foundry entities, as well as a general Aristotelian definition of *Nudge*: “A function of an environment (or part thereof) that is designed to influence behavior, in order to contribute to the influencee’s welfare or the society’s welfare, and that is easily resistible.” This classification distinguishes nudges, nudge-setting processes and nudge-activation processes, which is important to categorize correctly and disambiguate nudge-relevant data.

Further refinements could analyze more finely the processes related to nudge-setting and nudge-activation processes, following the classes introduced by BCIO [8]. The intention behind a nudge-setting process could also be analyzed in light of a theory of dispositions: since a nudge is a disposition, the intention to set up a nudge might be seen as a *pre-disposition*<sup>5</sup>, and the connection between the nudger’s intention to set up a nudge and his intention to influence behavior should be clarified. Other refinements may require first to clarify the ontological nature of central constructs in psychology (“heuristics”, “influence process”), philosophy (“rationality”), economics (“welfare”, “Econs”) and politics (“libertarian paternalism”). Future works should also define related behavioral interventions such as boosts [43].

## References

- [1] R. H. Thaler and C. R. Sunstein, *Nudge*. New Have, Connecticut: Yale University Press, 2008.
- [2] K. Van Ittersum and B. Wansink, Plate size and color suggestibility: The Delboeuf illusion’s bias on serving and eating behavior, *Journal of Consumer Research* 39 (2012) 215–228.
- [3] A. Barton, How tobacco health warnings can foster autonomy, *Public Health Ethics* 6:2 2013 207–219.
- [4] S. Benartzi, E. Peleg, and R. H. Thaler, Choice architecture and retirement saving plans, in *The behavioral foundations of public policy*, Princeton University Press, 2013, pp. 245–263.
- [5] C. Schubert, Green nudges: Do they work? Are they ethical?, *Ecological Economics* 132 (2017) 329–342.
- [6] B. Smith et al., The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration, *Nat Biotech* 25 (2007), 1251–1255. doi: 10.1038/nbt1346.
- [7] R. Arp, B. Smith, and A. D. Spear, *Building Ontologies with Basic Formal Ontology*, The MIT Press, 2015.
- [8] S. Michie et al., Representation of behaviour change interventions and their evaluation: Development of the Upper Level of the Behaviour Change Intervention Ontology, *Wellcome Open Research* 5 (2020).
- [9] J. Röhl and L. Jansen, Representing dispositions, *J. Biomed. Semant.* 2:S4 (2011). doi: 10.1186/2041-1480-2-S4-S4.
- [10] A. Barton, O. Grenier, L. Jansen, and J.-F. Ethier, The identity of dispositions, in: *Formal Ontology in Information Systems. Proceedings of the 10th International Conference (FOIS 2018)*, vol. 306, 2018, pp. 113–126.
- [11] A. D. Spear, W. Ceusters, and B. Smith, Functions in basic formal ontology, *Applied Ontology* 11 (2016) 103–128.

---

<sup>5</sup> I thank an anonymous reviewer for this suggestion.

- [12] S. Seppälä, A. Ruttenberg, and B. Smith, The functions of definitions in ontologies, in: *Formal Ontology in Information Systems. Proceedings of the 10th International Conference (FOIS 2016)*, 2016, pp. 37–51.
- [13] V. Berthet and B. Ouyard, Nudge: Towards a Consensus View?, *Psychology and Cognitive Science* 5 (2019) 1–5.
- [14] L. Congiu and I. Moscati, A Review of Nudges: Definitions, Justifications, Effectiveness, *Journal of Economics Survey* (2021), 1-26. <https://doi.org/10.1111/joes.12453>.
- [15] C. R. Sunstein, Misconceptions about nudges, SSRN 3033101, 2017.
- [16] P. G. Hansen, The definition of nudge and libertarian paternalism: Does the hand fit the glove?, *European Journal of Risk Regulation* 7 (2016) 155–174.
- [17] P. Mongin and M. Cozic, Rethinking nudge: not one but three concepts, *Behavioural Public Policy*, 2 (2018) 107–124.
- [18] R. Samuels, S. Stich, and M. Bishop, Ending the Rationality Wars: How to Make Disputes About Human Rationality Disappear., in *Common Sense, Reasoning and Rationality*, Oxford University Press, Oxford, 2002, pp. 236–268.
- [19] A. Barton and T. Grüne-Yanoff, From libertarian paternalism to nudging—and beyond, *Review of Philosophy and Psychology* 6 (2015) 341–359.
- [20] R. Münscher, M. Vetter, and T. Scheuerle, A review and taxonomy of choice architecture techniques, *Journal of Behavioral Decision Making* 29 (2016) 511–524.
- [21] P. L. Buttigieg, N. Morrison, B. Smith, C. J. Mungall, and S. E. Lewis, The environment ontology: contextualising biological and biomedical entities, *Journal of biomedical semantics* 4 (2013) 1–9.
- [22] L. Congiu and I. Moscati, Message and Environment: a framework for nudges and choice architecture, *Behavioural Public Policy* 4 (2020) 71–87.
- [23] C. Mills, The Heteronomy of Choice Architecture, *Rev.Phil.Psych.* 6 (2015) 495-509. doi: 10.1007/s13164-015-0242-7.
- [24] F. Guala and L. Mittone, A Political Justification of Nudging, *Rev.Phil.Psych.* 6 (2015) 385-395. doi: 10.1007/s13164-015-0241-8.
- [25] M. Nagatsu, Social Nudges: Their Mechanisms and Justification, *Rev.Phil.Psych.* 6 (2015) 481-494. doi: 10.1007/s13164-015-0245-4.
- [26] F. Toyoshima, A. Barton, and O. Grenier, Foundations for an ontology of belief, desire and intention, in: *Formal Ontology in Information Systems. Proceedings of the 11th International Conference (FOIS 2020)*, vol. 330, 2020, pp. 140–154.
- [27] A. Bandrowski et al., The ontology for biomedical investigations, *PloS one* 11:4 (2016) e0154556.
- [28] D. M. Hausman and B. Welch, Debate: To Nudge or Not to Nudge, *Journal of Political Philosophy* 18 (2010) 123–136.
- [29] P. G. Hansen and A. M. Jespersen, Nudge and the Manipulation of Choice: A Framework for the Responsible Use of the Nudge Approach to Behaviour Change in Public Policy, *Eur. J. Risk Reg.* 4 (2013) 3–28.
- [30] R. Rebonato, *Taking liberties: A critical examination of libertarian paternalism*, Palgrave Macmillan, New York, NY, 2012.
- [31] R. J. Arneson, Nudge and shove, *Social Theory and Practice*, 41 (2015), 668–691.
- [32] L. Bovens, The Ethics of Nudge, in: T. Grüne-Yanoff and S. O. Hansson (Eds.), *Preference Change*, Springer Netherlands, 2009, pp. 207–219.
- [33] Y. Saghai, Salvaging the concept of nudge, *Journal of medical ethics* 39 (2013) 487–493.
- [34] L. Jansen, Tendencies and other realizables in medical information sciences, *The Monist*, 90:4 (2007) 534–554.
- [35] A. Barton, A. Burgun, and R. Duvauferrier, Probability assignments to dispositions in ontologies, in: *Formal Ontology in Information Systems*, in: *Proceedings of the Seventh International Conference (FOIS 2012)*, vol. 239, 2012, pp. 3–14.
- [36] A. Barton, R. Rovetto, and R. Mizoguchi, Newtonian forces and causation: a dispositional account, in: *Formal Ontology in Information Systems*, *Proceedings of the 8th International Conference (FOIS 2014)*, vol. 267, 2014, pp. 157–170.
- [37] W. Hagman, D. Andersson, D. Västfjäll, and G. Tinghög, Public views on policies involving nudges, *Review of philosophy and psychology* 6 (2015) 439–453.
- [38] W. DeJong et al., A multisite randomized trial of social norms marketing campaigns to reduce

- college student drinking, *Journal of studies on alcohol*, 67 (2006) 868–879.
- [39] A. Barton, L. Jansen, and J.-F. Ethier, A taxonomy of disposition-parthood, in: *Proceedings of the Joint Ontology Workshops 2017 (JOWO 2017)*, vol. 2050, Bolzano, Italy, 2018, pp. 1–10.
- [40] R. Baldwin, From regulation to behaviour change: giving nudge the third degree, *The Modern Law Review*, 77 (2014) 831–857.
- [41] C. R. Sunstein, People prefer system 2 nudges (kind of), *Duke LJ* 66 (2016).
- [42] Y. Lin, M. Osman, and R. Ashcroft, Nudge: concept, effectiveness, and ethics, *Basic and Applied Social Psychology* 39 (2017) 293–306.
- [43] R. Hertwig and T. Grüne-Yanoff, Nudging and boosting: Steering or empowering good decisions, *Perspectives on Psychological Science* 12 (2017) 973–986.