

## Response letter to reviewer #1's comments

### Comment 1:

*"The paper's main contribution is as a review of literature on nudges (soft paternalistic policies based on behavioral economics insight), organized through the lens of the dual process model written to be friendly to policy makers and interested voters who are grappling with the idea of nudges."*

### Reply to comment 1:

Thank you very much. Indeed, it was our aim to review the existing literature on nudges and sharpen the definitional scope by explicitly focusing on insights from cognitive sciences. By doing so, we aim to bring clarity to the existing literature where the scope of nudges as a policy instrument is often plurivalent and therefore includes a number of policy instruments where the classification is not clear.

### Comment 2:

It essentially is a lit review, combining the insights of Thaler and Sunstein's Nudge on soft paternalism using the framework presented Kahneman's Thinking Fast/Thinking Slow on the dual process model. While written in the style of a research paper, it feels very much like a recapitulation of those two popular books.

### Reply to comment 2:

It is correct that there is a basic overlap of contents between those two popular books and our manuscript as those books address the topic of nudges respectively the dual process theory. However, there are substantial differences regarding the analytical scope and the scientific depth of our analysis.

The definition of a nudge given by Thaler and Sunstein in their book mentions some important characteristics (being effective by varying the choice architecture, formally non-restrictive); however, it lacks a more in-depth explanation why individuals "suffer" from so-called cognitive biases mentioned in the book (e.g., anchoring, status quo bias) and how a more detailed assessment of the underlying behavioral principles helps us to classify the various existing, non-restrictive policy tools (e.g. monetary incentives, information campaigns, and nudges).

Regarding Kahneman's popular book, there are also some major differences. In his book entitled "Thinking, Fast and Slow", Kahneman explains the two systems (intuitive/fast vs. deliberative/ slower) that determine the way we think. Then, he applies this systematic ap-

proach in order to describe behavioral phenomena that are labelled as “biases” from a rational perspective. While our essay relates to the basic topic of this book (i.e., dual system/process theory), it goes a step beyond by including further insights from cognitive sciences on the boundaries and interactions between the two processes. It also provides an overview on the frequency of different types of decision process which is important to conceptualize nudges (see page 4 et seq. of the manuscript). Hence, the scope and scientific depth of the essay is far beyond of those popular books.

In order to improve our essay according to your suggestion, we will include explanations on distinctions between those popular books and our essay in the introduction of the sections “2. *Human behaviour from the perspective of cognitive and social psychology*” and “3. *What is nudging and how does it work?*” where we will indicate that both topics are also dealt with in popular books and will explain in detail to what extent our analysis deepens the frameworks of those popular books.

**Comment 3:**

*“It attempts to sharpen our definition of nudges by focusing on policies that affect behavior through Type I (instinctual processes) rather than policies that affect behavior via information provision that operates through Type II (cognitive processes). I don’t think it goes very far beyond what could be learned from reading those books, however it summarizes the literatures in a coherent way, and seems consistent in style and spirit to other summary papers published in this outlet.”*

**Reply to comment 3:**

As partly discussed in the response to comment no. 2, the review in our essay provides an analysis of nudges that differs from the framework presented in the Thaler and Sunstein book. In contrast to Thaler and Sunstein, we do *not* follow a casuistic approach to describe nudges and elaborate on their effectivity. Instead, we develop a definition which is grounded in the cognitive sciences literature on decision making and dual process theory, and use this definition to provide a classification of nudges. This again allows to separate nudges in the context of policy instruments from other non-restrictive (“soft”) policy measures.

The missing definitional link of nudges to underlying principles from cognitive sciences has blurred the definitional clarity and scope of the nudging concept in the literature, and led to the emergence of different interpretations depending on the research question and/ or area of application concerned (e.g., Ölander and Thøgersen 2014; for a discussion of ‘mistaken nudges’ see e.g. Selinger and Whyte 2011, Hausman and Welch 2010). With a clear reference to intuitive cognitive processes we are able to distinguish nudges from other non-

restrictive policy instruments: *“In order to clarify this issue, we argue that behavioural interventions that are effected automatically and intuitively (process 1) and without restricting the choices available are at the heart of the nudging paradigm (Grüne-Yanoff and Hertwig 2015).”* (see page 5-6).

However, findings from cognitive psychology (e.g. Evans and Stanovich 2013) also indicate that this definitional approach has to be extended as *“in many cases process 1 output serves as a starting point for an informed decision and thus has only an indirect impact on the final decision”* (see page 6 of the manuscript). This means there are (policy) measures that target reflective decision-making processes but reach to influence those processes through an underlying intuitive and automatic stimulus (e.g. so-called ‘social’ nudges which are based on peer effects and related stimuli). In response, we extend the definition and follow the line of argument suggested by Hansen and Jespersen (2013) so that we classify nudges as interventions that always aim at influencing the intuitive cognitive process (process 1) and evoke a behavioral change either directly (“pure” type 1 nudge) or indirectly through cognitive process 2 (type 2 nudge) (see pages 6-8 and figure 1).

In the text, we exemplify the first sub-category of nudges with the help of the ambient-orb device. This device indicates real-time energy use by means of colored light signals which can be interpreted as

- I.) information that is processed through a deliberative thinking process
- II.) or may also evoke an unconscious response as a result of an intuitive association with a displayed color (see e.g. Selinger and Whyte 2011).

Here, the cognitive science perspective would lead to a classification of the ambient orb device as a type 1 nudge if the behavioral change is invoked by a stimulus in the domain of the intuitive thinking process (second interpretation) while the first interpretation would imply that the ambient orb device is an information tool and therefore not a nudge. However, we have to admit that the example given should be reconsidered as the starting point where this example is drawn from is the discussion on so-called “fuzzy nudges”.

In order to better clarify the distinction between type I and type II nudges, we propose to include in a revise-and-resubmit version further examples. We will develop a table that contrasts type 1 and type 2 nudges so that the reader can grasp the difference between both categories.

#### **Comment 4:**

*“It is written clearly, although given its brevity, it sometimes feels selective, picking and choosing a few things to go deep in, but ignoring others that I would think could have been*

*given equal weight. For example, why focus on the rather obscure “query theory,” when there are many other theories that are more commonly cited when talking about defaults.”*

**Reply to comment 4:**

In the article we focus on the implications of the Dual Process models for human behavior and describe several possible reasons for the effectiveness of technical defaults from this particular perspective. We purposefully leave out legal economic default theories that do not fit this specific cognitive science framework of our paper. The above mentioned reasons can be roughly divided into two groups: (1) limited cognitive capacity to make a decision which can result from imposed circumstances or a “rational” reflected decision to optimize limited resources (here time), (2) a range of *cognitive biases* such as implicit recommendation, status quo bias, loss aversion and omission bias described in the paper. The latter group, closely connected to the intuitive and heuristic driven process behind the nudging approach, is the main focus of our work. The query theory, that has been cited in our survey for the sake of comprehensiveness, falls clearly into the second group as the default option is often perceived as an anchor in reference to which all other options are compared and evaluated (another cognitive bias). Thus, it needs at least to be mentioned in line with our chosen approach.

**Comment 5:**

*“The second half engages in some speculation. For example, it extends a bit into how behavioral nudges interact with existing policies. Here it recapitulates some of the discussion found in another popular book (Congdon, Kling and Mullainathan) but provides little concrete research”*

**Reply to comment 5:**

Thank you for the reference to the book Policy and Choice by Congdon, Kling and Mullainathan – we will include the citation at the appropriate place.

Indeed, in their book the authors also handle the topic of policy interventions encompassing both, well-established economic tools like taxes and behaviorally-informed interventions like nudges. In an excellent way the authors make it very clear that understanding of the behavioral aspects can give us a full picture of the market structure and performance. However, as in the case of Thaler and Sunstein book, the definition of nudges leaves some room for an individual interpretation. This can be explained by the lack of solid psychological background that is needed to clearly differentiate nudges from other interventions and understand the mechanism of their work. In addition, our assessment of the state of art provides an up-date on some aspects mentioned in the book – e.g. the notion of nudges.

**Comment 6:**

*“There is related research on the interaction between hard regulations and behavioral factors that the authors allude to but do not discuss, such as in Bao and Ho and Bowles and Polania-Reyes which the authors cite; or the substantial literature on trust and incentives (Falk and Kosfeld, 2006), or the substantial literature on incentives and pro-social behavior related to Benabou and Tirole (2006).”*

**Reply to comment 6:**

In the manuscript, we start our discussion on the interaction between “traditional” policy instruments (e.g. bans, monetary incentives) and nudges by bringing together insights from cognitive processing and human decision making with the mode of operation of these traditional tools in order to develop an analytical framework that is consistent with the way we approach nudges. In this context, we also briefly discuss possible consequences of what can happen when informational effects of monetary incentives are accompanied by behavioral patterns that operate through the intuitive and affective process (e.g. salience, the importance of reference points, or social norms) and come to the conclusion that the large number of empirical and experimental studies that include examples where price instruments tend to be not as effective as expected can be explained by this interaction intuitive (process 1) and rational (process 2) thinking (see page 21 of the manuscript for the case of monetary incentives).

The literature proposed by the reviewer, in difference, ties to the theory of interactions of monetary incentives and social behavioral patterns. The theoretical model of Benabou and Tirole (2006) is in line with many other important works on the effects of pro-social behavior of economic agents in a decision situation (see also Rabin 1993; Fehr and Schmidt 1999; Bolton and Ockenfels 2000). Those models have in our view an important impact on the discipline of economics as they contributed significantly to the discussion on behavioral assumptions of (micro-)economic modeling and succeeded to explain results from experimental and empirical research that could not be explained with the homo economicus framework.

Linking this literature to our approach could serve as a starting point in order to assess effects of nudges in a mixed use. However, as mentioned before, our approach is deeply rooted in cognitive sciences’ dual process theory and in policy tools that address the duality of intuitive and deliberate thinking and work explicitly with the influence of intuitive thinking on human decision making. The research strand that is mentioned by the reviewer focuses rather on incentives which work through a rather singular channel (i.e. the deliberate thinking

process) and extend the perspective on these types of policy interventions in order to explain why in some studies their effect was not as expected.

An analysis of possible interaction effects of nudges has to begin from another starting point, i.e. the intuitive thinking process and the duality of both, intuitive and deliberate thinking processes. Existing work that starts from this perspective indicates that a mixed use of policy tools is useful as such mixed strategies allow to address simultaneously individuals that in a specific setting decide intuitively as well as individuals that in the same setting decide through a deliberative process (see e.g. Goldin and Lawson 2015; see also page 21-22 of the manuscript). However, the intuitive stimulus (particularly intuitive social stimuli) has to be chosen with care as the experimental and empirical findings on monetary incentives shows.

The literature suggested by the anonymous commentator on the journal's website (Dolan and Galizzi 2015) deepens this question and discusses promoting spillover, permitting as well as purging spillovers. In this regard, several strands of research have looked at the framework conditions that determine the occurrence of these spillovers. Among these research strands are: questions regarding the relative costs of behavior, trade-offs between different motives, and the cognitive mindset during interacting behaviors where the latter is the focus of our paper (Dolan and Galizzi 2015)

Here, our general conclusion is that in order to inform policy, scientists and policy makers have to capture all the ripples from one behavior to the next when a single policy measure or mixed strategy is elaborated. In the context of nudges, there is to date relatively little systematic research on this key point (see Dolan and Galizzi 2015; page 22 of the manuscript). Hence, it is difficult to draw a general conclusion on effects that occur in a mixed use of nudges.

In order to deepen the analysis, we propose to include the discussion on different research strands that help to assess possible spillovers when revising our manuscript.

**Comment 7:**

*“There is some nice novel speculation here about what the authors call “double nudges” and how multiple nudges might interact with each other, but this is pure speculation and suggestions for future research.”*

**Reply to comment 7:**

In our assessment of the status quo we wanted not only to summarize the available results in the field of nudging research, but also point at and, where possible, clarify open points. One of such points is the under-researched topic of the interaction between different types of nudges which we call a “double nudge”. Taking into account a scarce number of papers referring to this interesting research question, the only reasonable way to tackle this topic was

an interdisciplinary approach using the well-established psychological theory which is also in line with the scope and focus of our paper. This enabled us to pose a couple of interesting hypotheses, which are clearly stated and which – as we noted – require further empirical verification. By doing so, we hope to evoke interest in those specific topics and encourage further empirical research in this domain in order to validate or reject our hypotheses based on the insights from cognitive sciences.

In order to improve the structure of the article, we intend to move the part referring to the notion of double nudges to the last section of the paper which will now include also future outlook.

**Comment 8:**

*“It also calls to attention the understudied importance of habit, although there has been some relevant recent work by say Landry or Karlan, and the foundational work by Becker and Murphy that could have grounded this section better.”*

**Reply to comment 8:**

The seminal paper of Becker and Murphy (1988) handling the topic of rational addiction as well as the relatively new extension of their model by Landry (2014) provide significant contribution to the economic theory of habit (and addiction). The work by Landry (2014) incorporates some important insights from the psychology of habit formation such as the necessity of a thoughtful (reflected) decision to engage in a particular behaviour or the role of the external cues in reinforcing the persistence stage which are points of interception with our paper.

However, we have to note that those papers suggested by the referee rely on a totally different framework with the utility maximizing actors that is incompatible with our approach based on the cognitive and social psychological theories. In particular, actions guided by the intuitive and automatic process 1, which is an important pillar of our analysis, can result in choices inconsistent with the rational choice models.

**Comment 9:**

*“Finally, by focusing on the dual process model as the organizing principle for this review, the paper neglects much discussion on the single largest part of the nudge literature, peer effects and peer information. It cites many related papers like Alcott and Rogers, and alludes to much of the work by Beshears et al. on such peer nudges, but it seems like a flaw that its conceptual framework says little about peer effects generally.”*

**Reply to comment 9**

Peer and neighborhood effects are important parts of the social economics explaining the influence of a group on an individual (Becker and Murphy 2000). However, standard economic methods, in contrast to social psychology, do not account for the differences in the cognitive processing of specific social norms that influence an individual behaviour and that are crucial to our classification of the policy instruments. For this reason we deliberately omitted socio-economic theories of peer effects and focused on the purely psychological ones.

In the paragraph about the effectiveness of the double nudges we referred to the work of Cialdini (2003) indicating that descriptive norms (what most people do) are rather intuitive (and thus likely to be dominated by the automatic process 1 as a heuristic or cue what to do based on the observed behaviour of others). On the other hand, the injunctive norm (what is believed to be correct in the society, operating through the social sanctions and peer pressure) was proven to require more deliberate processing (and thus can be expected to be governed by the deliberate process 2).

Following the recommendation of the referee we will try to give more weight to this aspect and add additional references supporting that claim, e.g. the study of Ohtomo and Hirose 2007.

**Comment 10:**

*“Overall this paper offers another update primarily for policy makers on the current literature about nudges. Here, it joins the popular books mentioned above, or the various other JEP/JEL style papers like the Madrian paper it cites. Its main novelty is its use of the dual process model to frame the discussion, and its speculation about the role of habit and interactions between policies.”*

**Reply to comment 10:**

This summary statement reflects perfectly the intentions and the innovations geared at with our survey paper. However, we believe that this review can be also helpful to the economists dealing with the question of policy instruments and behavioral economics – it helps to systemize the knowledge on the topic of nudges and also explains the psychological mechanism underlying nudging interventions which differs substantially from other “soft” instruments like pure information provision or moral suasion, to name only a few.