Szászi Barnabás Imre

Investigating human decision-making: from processes to applications

Doctoral School of Psychology,
Head of the School: Dr. Zsolt Demetrovics, professor, Eötvös Loránd University

Behavioral Psychology Program
Head of the Programme: Dr. Anna Veres-Székely, professor, Eötvös Loránd University

Supervisor
Dr. Balázs Aczél, associate professor, Eötvös Loránd University

Committee members:
Dr. Klára Faragó, professor emerita, Eötvös Loránd University, President
Dr. Emese Józsa, assistant professor, Eötvös Loránd University, Secretary
Dr. Richárd Szántó, associate professor, Corvinus University of Budapest, Opponent
Dr. László Mérő, professor, Eötvös Loránd University of Budapest, Opponent
Dr. Péter Simor, assistant professor, Eötvös Loránd Tudományegyetem, Member
Dr. Dezső Németh, professor, Eötvös Loránd Tudományegyetem, Member
Dr. Péter Pajkossy, assistant professor, Budapesti Műszaki Egyetem, Member

Budapest, 2018
List of publications forming the basis of the dissertation


Introduction

The investigation of factors influencing human decision-making and the exploration of their underlying cognitive processes form the basis of this dissertation. In the first three studies, I examine how two factors, the first quick thoughts coming to mind and the action dynamics of thinking influence individuals’ decision-making. By tracking and manipulating the underlying cognitive processes, I tested the predictions of the default-interventionist and the embodied choice theories. One goal of the decision-making research is to apply the results of the fundamental research in applied setting. Accordingly, extending the scope of the dissertation, I pose a more general question in the final study: how scientific practices should be changed if as a field, we want to achieve more effective evidence accumulation from the studies applying choice architecture interventions.

In the last decades, hundreds of factors influencing decision processes have been explored. The paradigmatic approach of the field, the heuristics and biases (HB) research program, investigated the different ways in which some of these influencing factors can cause systematical deviations from some normative standard or optimum (e.g., preference reversals, hindsight bias, framing, outcome bias, planning fallacy, sunk cost effect). Numerous competing explanations were developed on why such biases occur. Although there is no single theory which could account for the why's behind all of the biases, there have been some attempts in the recent years to create overarching models providing explanations at least for some of them.

One such theory is the default-interventionist theory, which assumes that many of these biases happen because individuals produce an incorrect quick thought (heuristic response) and fail to override this first response with further deliberation (Evans, 2007; Kahneman & Frederick, 2002; Stanovich, 2009). Although there are many different kinds of default-interventionist theories, they all share two distinctive features regarding the temporal pattern behind decision biases: an initial incorrect response and the (lack of) intervening deliberative process. However, in the recent years, many methodological concerns have been raised (e.g., Aczel, Bago, Szollosi, Foldes, & Lukacs, 2015) and contradictory findings have been found (Thompson & Johnson, 2014) resulting in the loss of confidence in some of the building pillars of the default-interventionist theory. In Study 1 and Study 2, my colleagues and I examine a core pillar of the default-interventionist theory using
the Cognitive Reflection Task (CRT, Frederick, 2005) and the denominator neglect task (Bonner & Newell, 2010): how the first intuitions and the lack of intervening analytic thinking influence individuals’ decisions.

As discussed, Study 3 investigates a different influencing factor. It focuses on the embodied choice theory which discusses how bodily movements (action dynamics) can influence decisions. The logic of the embodied cognition theories theory follows: if a concept is activated in the mind, it also activates the constituent sensory-motor states (Caramazza, Anzellotti, Strnad, & Lingnau; 2014). Conversely, if a sensory-motor state is activated (embodied) it leads to the activation of related concepts in the mind and even changes the evaluation of related concepts. Previous experiments (Burk, Ingram, Franklin, Shadlen & Wolpert, 2014; Lepora & Pezzulo, 2015; Resulaj, Kiani, Wolpert & Shadlen, 2009) did not allow researchers to disentangle which of the two possible reasons are responsible for the influence of bodily movements on the decisions: do they influence decisions because they change the cost of choosing one of the response options or because they change the value of the response options. In Study 3, my colleagues and I aim to disentangle the influence of these two effects.

Based on the insights of the studies examining the key effects influencing decisions, several theory-based behavioral interventions have been developed in order to support important societal goals. Such behavioral interventions aiming to influence individuals’ decisions were applied by government agencies in more than 40 countries having a significant positive impact on millions of people’s lives (e.g., Hansen, Skov, & Skov, 2016). However, due to methodological concerns first raised by the broader field of general behavioral interventions (e.g., Michie et al., 2016) it has been suggested that the accumulation of the scientific knowledge on when and why these interventions work is too slow. Accordingly, in Study 4, we intend to figure out how the field of nudging could more effectively accumulate knowledge on when and why nudges work and which nudge to use in a given situation to reach important policy goals. In order to identify the most important obstacles that can hinder the evidence accumulation of the field, we provide an overview of the characteristics of the interventions and experimental designs employed in the nudge literature. Additionally, we aim to convey recommendations for future studies.
Study 1

The cognitive reflection test revisited: exploring the ways individuals solve the test

**Aims:** In Study 1, we aimed to explore the strategies used and the abilities employed when individuals solve the cognitive reflection test (CRT), the most widely used measure of individuals’ propensity not to override the first answer that comes to mind. As this tendency is thought to be a crucial cause behind many failures in reasoning, the goal of Study 1 was twofold: first, to explore the role of first thoughts in biased decisions, second, to contribute to the deeper understanding of the “whys and whens” of decision errors in heuristics and biases tasks by exploring the role of individual differences in the applied strategies and in the CRT performance.

**Methods:** 219 students (75% female, $M = 22.04$ years, $SD = 2.28$) participated in our study. The study consisted of an offline and an online session. For the offline session, participants were invited to the lab to participate in a personal interview. First, they were informed that the session would be recorded and later analyzed. This was followed by the detailed verbal instruction of the protocol and a warm-up. After that, participants were asked to solve the three items of the CRT in the standard order whilst thinking aloud. During the online sessions, participants completed several reflectivity (such as the Rational-Experiential Inventory, REI; Pacini & Epstein, 1999) and ability measures (Berlin Numeracy Test, BNT; Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012) in a fixed order.

**Results:**

1) First, we found that both numeracy (BNT, $b = .39, 95\% \ CI [0.29, 0.48], t = 8.22, p < .001$) and reflectivity and (REI, $b = 0.02, 95\% \ CI [0.01, 0.03], t = 4.16, p < .001$), accounts for the performance on the CRT $F(2,203) = 48.09, p < .001, adj. R^2 = .32$.

2) Protocol analysis was conducted: Two raters, blind to our hypotheses, categorized the verbal reports using the following coding system. The correct answers were classified into the ‘Correct start’, or the ‘Incorrect start’ categories. The incorrect responses were grouped as ‘Reflective’ or as ‘Non-reflective’.

   a. The protocol analysis of the correct answers suggests that the participants performed a ‘Correct-start’ in 124 cases (77%) and showed an ‘Incorrect start’ pattern only in 37 cases (23%). Furthermore, we analysis of the reaction time data
indicated that the reasoners did not suppress their incorrect first thoughts before articulating their answer \((B_{H(0, 1.63)} = 0.28)\).

b. The protocol analysis of the incorrect answers aimed to explore whether there are people who check the first answer that comes to their mind but still fail to solve the task. The data suggest that in 142 of the 361 cases (39%) people engaged in some kind of reflective behavior after reporting their first answer, while in 219 cases (61%) people accepted the first answer that they reported without any further deliberation.

3) Our data provided insensitive results \((B_{H(0, 0.45)} = 0.62)\) regarding the hypothesis that more numerate individuals start their thinking with correct strategies or have correct intuitions on the CRT more often than their low numeracy counterparts. Similarly, Our results were insensitive \((B_{H(0, 0.03)} = 0.80)\) regarding the hypothesis that people in the ‘Reflective’ group score higher on the REI scale than the members of the ‘Non-reflective’ group.

**Discussion:** The protocol analysis revealed that there are several ways people solve or fail the test. Importantly, 77% of the cases in which reasoners gave the correct final answer in our protocol analysis, they started their response with the correct answer or with a line of thought which led to the correct answer. We also found that 39% of the incorrect responders reflected on their first response. The findings indicate that the suppression of the first answer may not be the only crucial feature of reflectivity in the CRT and that the lack of relevant knowledge is a prominent cause of the reasoning errors. Additionally, we confirmed that the CRT is a multi-faceted construct: both numeracy and reflectivity account for performance.
Study 2

Investigating the dynamics of individual differences in the denominator neglect task

Aims: Most decision-making models describing individual differences in heuristics and biases tasks build on the assumption that reasoners produce a first incorrect answer in a quick, automatic way which they may or may not override later and that the advantage of high capacity reasoners arises from this late correction mechanism. The goal of Study 2 was twofold. First, we aimed to investigate the assumption that reasoners first produce incorrect answers in HB tasks. Accordingly, we explored the proportion of trials in which individuals moved their mouse initially towards the correct response. Second, we investigated why higher capacity people give more normative answers. Specifically, we tested three explanations: Higher capacity people (1) have a higher likelihood for initially correct answers, (2) are more likely to stay with their initial answer when it is correct, and (3) are more likely to change their mind when their initial answer is incorrect.

Methods: The experiments consisted of two sessions, an offline and an online session. For the offline session, participants were invited in groups of 15-20 to a computer test room. Participants were presented with 54 of a computerized version of the denominator neglect tasks while the mouse movements were recorded. In the denominator neglect task, participants are presented with two ratios and are asked to choose the larger one (Bonner & Newell, 2010).

To assess the dynamics of thinking in the denominator neglect tasks, we developed a mouse-tracking analysis using the areas of interest (AOI) technique (Figure 1). The main idea behind this technique is that one can explore a reasoner’s first and subsequent choice commitments by creating AOIs surrounding the choice options and analyzing the order in which the AOIs were visited by the mouse cursor in each trial. In the current study, we used the reasoner’s initial commitment (i.e., first AOI around one of the choice options visited by the participant’s mouse cursor) as a proxy for the participant’s first answer. If this first commitment differed from the individual’s final answer, we classified the trial as a Change of Mind (CoM) trial.
Figure 1. Figure 1A shows an exemplary ratio pair along with the borders of the corresponding areas of interests (AOIs) and the starting position of the mouse cursor. Figure 1B, 1C and 1D show three exemplary mouse trajectories.

After the offline session, participants were sent an email containing the information about the online session. Here, they were asked to fill out an online test package containing the cognitive capacity measures (IQ test - Kovacs & Temesvári, 2016; BNT - Cokely et al., 2012).

Results: The current paper contains two experiments: a mouse-tracking experiment and its replication. As the methods and the analyses were identical for both experiments, we report them conjointly. We recruited 109 and 143 participants in experiment 1 and experiment 2 respectively.

We consistently observed across two experiments that individuals’ mouse cursor was often drawn first toward the correct answer (Table 1 summarizes the results).
Table 1

Percent of Trials (in the incongruent trials) per Experiment Classified Based on the Correctness of the Initial and Final Response

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Correct final response</th>
<th>Incorrect final response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td>Correct initial response 42%</td>
<td>Incorrect initial response 42%</td>
</tr>
<tr>
<td></td>
<td>Incorrect initial response 3%</td>
<td>Incorrect initial response 13%</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>Correct initial response 40%</td>
<td>Incorrect initial response 45%</td>
</tr>
<tr>
<td></td>
<td>Incorrect initial response 3%</td>
<td>Incorrect initial response 11%</td>
</tr>
</tbody>
</table>

We also found evidence that cognitive capacity is one important predictor of the individual differences and this capacity-normativity relationship arises from late and not from early processes: cognitive capacity did not predict the normativity of the first answers but it was associated with more changes of minds after incorrect first answers (Table 2).

Table 2

Results of Generalized Linear Mixed Models Using the Cognitive Capacity Score to Predict the Correctness of the Initial and Final Response in the incongruent condition in Experiment 1 and 2

<table>
<thead>
<tr>
<th>Predicted variable</th>
<th>Experiment</th>
<th>OR</th>
<th>95% CI</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correctness of final response</td>
<td>1</td>
<td>1.43</td>
<td>[1.13, 1.81]</td>
<td>3.01</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.58</td>
<td>[1.30, 1.93]</td>
<td>4.62</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Correctness of initial response</td>
<td>1</td>
<td>1.05</td>
<td>[0.93, 1.18]</td>
<td>0.80</td>
<td>.424</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.98</td>
<td>[0.90, 1.07]</td>
<td>-0.45</td>
<td>.656</td>
</tr>
</tbody>
</table>

Note. All models are generalized linear mixed models with a binomial link function.
Interestingly, we observed that high capacity individuals make less incorrect changes and more correct changes after the first answers and this effect also significantly contributed to the emergence of the capacity normativity relationship.

Table 3

Results of Generalized Linear Mixed Models Using the Cognitive Capacity Score to Predict the Occurrence of a Change of Mind Depending on the Correctness of the Initial Answer in the Incongruent Condition in Experiment 1 and 2

<table>
<thead>
<tr>
<th>Initial answer</th>
<th>Experiment</th>
<th>OR</th>
<th>95% CI</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect</td>
<td>1</td>
<td>1.45</td>
<td>[1.11, 1.90]</td>
<td>2.75</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.77</td>
<td>[1.42, 2.20]</td>
<td>5.06</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Correct</td>
<td>1</td>
<td>0.67</td>
<td>[0.47, 0.95]</td>
<td>-2.28</td>
<td>.023</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.70</td>
<td>[0.53, 0.93]</td>
<td>-2.43</td>
<td>.015</td>
</tr>
</tbody>
</table>

Note. We used generalized linear mixed models with a binomial link function.

Discussion: In sum, our study suggest that reasoners may not always produce an incorrect answer first in heuristics and biases task. Furthermore, we observed that compared to low capacity reasoners, high capacity individuals revise their first answer more frequently if it is incorrect and make fewer changes if it is correct. However, we did not find evidence that high capacity individuals produce correct initial answers more frequently. Consistent with the predictions of previous default-interventionist decision-making models, these results suggest that in the denominator neglect task the capacity-normativity relationship arises after the initial response is formulated. Finally, our study showcases how mouse-trajectory analysis can be utilized to investigate individual differences in decision-making.
Study 3


Aims: In Study 3 focused on the embodied choice theory which discusses how bodily movements (action dynamics) can influence decisions. According to the hypothesis of embodied choice, the decision-making process is bidirectional as action dynamics exert their backward influence on decision processes through changing the cost and value of the potential options. This influence takes place as moving towards one option increases the commitment to and, therefore, the likelihood of choosing that option. This commitment effect can be the result of either (a) the continuous act of getting closer to this option, or (b) the increased movement cost associated with changing the movement direction to select a different option. In study 3 we aimed disentangle the potential influence of these two factors.

Methods: We developed the Guided Movement Task (Figure 4), a choice task designed to bias participant’s computer-mouse movements by constraining the allowed movement space by a corridor. Using this task, we created different conditions in which the participants’ mouse-cursor, after being guided towards one of the options, either had equal or unequal distances to the choice options. By this manipulation, we could test whether the continuous act of getting closer to an option in itself is sufficient to influence people’s decisions – a claim of “strong embodiment”.

The current paper contains two experiments. Both experiments applied the GMT task but used different stimuli to test our hypotheses. For the first experiment, 180 portrayal pictures (90 male, 90 female) were gathered from the Chicago Face Database, and people asked “Which one of the following faces do you find more attractive?”. In the second experiment, the stimuli consisted of 43 gamble pairs created by Koop and Johnson (2013, Study 2). In each trial, people were asked the following question: “Of the following two options, which one would you prefer to choose?”. We recruited 115 university students (98 females; $M = 22.36$ years, $SD = 2.04$ years) in experiment 1 and 194 students (150 females; $M = 21.82$ years, $SD = 1.80$ years) in experiment 2.
Figure 4. Summary of the experimental manipulations in Experiment 1. The distances to the options after the initial movement are unequal in panel A (shorter distance to the right option) and B (shorter distance to the left option), and equal in panel C and D. The initial movement is directed towards the right option in panel A and C, and directed towards the left option in panel B and D.

**Results:** In two experiments, we found that the likelihood of choosing an option only increased when the distances between the two options were unequal after the initial movement, but not when they were equal.

In experiment 1, the effect of direction was not significant in the baseline condition when the distances to the options were equal, $OR = 0.95$, 95% CI [0.85, 1.07], $z = -0.82$, $p = .41$. The effect of direction was significant in an additional model where the unequal distance, direction left condition served as the baseline, $OR = 2.17$, 95% CI [1.93, 2.44], $z = 13.03$, $p < .001$. 
Figure 5. The probability of right-option choices in Experiment 1 depending on the initial movement direction (towards the left vs. right option) and the distances to the options after the initial movement (equal vs. unequal)

In experiment 2, replicating the findings of Experiment 1, we found no significant effect of direction on the proportion of right-side choice in the equal distance conditions, OR = 0.94, 95% CI [0.83, 1.07], z = -0.94, p = .35. As in Experiment 1, the effect of direction was significant when using the unequal distance, direction left condition as the baseline, OR = 1.46, 95% CI [1.29, 1.66], z = 5.76, p < .001.

Discussions: Across both experiments, the initial movement towards an option did not bias participants’ choices when the distances to the options were equal. This bias only occurred when the initial movement also led to an increase in distance to the other option. In other words, these results suggest that moving towards an option in itself, does not increase commitment, and it hence does not increase the preference of the option. These results disagree with the hypothesis that action execution is an inherent part of the decision-making process.
A Systematic Scoping Review of the Choice Architecture Movement: Towards Understanding When and Why Nudges Work

**Aims:** In the endeavor of helping people make beneficial decisions for society and for themselves, the implementation of *choice architecture* interventions became a promising approach. However, due to methodological concerns first raised by the broader field of general behavioral interventions (e.g., Michie et al., 2016) it has been suggested that the accumulation of the scientific knowledge on when and why these interventions work is too slow. In study 4, first, we attempt to give an overview of the characteristics of the interventions and experimental designs employed in the nudge literature. Second, we intend to identify the most important obstacles that can hinder the evidence accumulation of the field and we aim to reveal whether the deficiencies found in specific domains characterize the field. Third, based on this overview and the insights of related disciplines, we will convey several recommendations for future studies.

**Methods:** We conducted a systematic literature search in three academic databases (Scopus, PubMed, and Pubpsych) with a predetermined search strategy in order to identify all the choice empirical architecture research after 2007. After the literature search, a screening procedure was conducted following predetermined criteria (Table 4). From the final list of target articles, one of the authors extracted the data and another author reviewed it.

<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Full-text papers.</td>
</tr>
<tr>
<td>(2) The paper’s language was English.</td>
</tr>
<tr>
<td>(3) The paper was published in a peer-reviewed journal.</td>
</tr>
<tr>
<td>(4) The studies in the paper empirically investigated one or more behavioral intervention technique that was attributed as a nudge or was connected to the choice architecture literature by the original authors.</td>
</tr>
<tr>
<td>(5) The studies in the paper had behavioral outcome variables (not preferences or attitudes).</td>
</tr>
</tbody>
</table>
Exclusion

(1) Review articles, conference abstracts, and conference papers.

(2) The studies in the paper applied interventions that restrict the freedom of choice of the target population, included significant economic incentives or used education, complex decision support systems or consultation as a nudge.

Results: 2670 entries were identified through the search in the three academic databases. After the screening of the articles, our final list contained 116 empirical articles with 156 studies and more than 422 tested interventions (See Figure 6).

We found that 74% of the studies were mainly motivated to assess the effectiveness of the interventions in the concrete settings, while only 24% of the studies focused on the exploration of moderators or underlying processes. We also observed that only 7% of the studies applied power analysis; no study in our database was preregistered; the used intervention nomenclatures are non-exhaustive and often have overlapping categories; and that no reporting guidelines have been used.
**Discussion:** Building on these observations and proposed solutions from other fields, we provide directly applicable recommendations to the field, which we think that could support the evidence accumulation on why and when nudges work.

- We urge the further development of the choice architecture taxonomies since without a comprehensive nomenclature with clearly labeled, non-overlapping and non-redundant categories. We suggest that the nudge and behavioral intervention taxonomies should be integrated. We also suggest adopting the taxonomy of the mode of deliveries.

- We also recommend the usage of reporting guidelines. Along many others (e.g., Munafo et al., 2017; Gonzales & Cunningham, 2015), we also recommend the usage of the public preregistration systems.

- We suggest that more studies should focus on the exploration of moderators or underlying processes.

**General discussion**

The present dissertation focused on the factors influencing and the cognitive processes underlying human decision-making. In the first two studies, we explored how individuals solve the Cognitive Reflection Test and the denominator neglect task. By tracking the thought processes, we tested the time-course assumptions of the default-interventionist theory, a central decision-making theory discussing the influencing role of the first thoughts and the lack of intervening analytic thinking behind many reasoning failure. The results of Study 1 and 2 suggested that the default-interventionist model needs some refinement. In Study 3, we created a novel task (the Guided Movement Task) and investigated the process by which action dynamics influence individuals’ decisions. The findings provided evidence that the movements influence the cost of the change but not the value of the options. In the last study, we conducted a systematic review of choice architecture interventions. These interventions use the insights from the psychology literature on the factors influencing decision-making. Based on the results of the systematic scoping review, we formulated several recommendations on how scientific practices should be improved in order to enhance the process of evidence accumulation on when and why nudges work and on which nudge to use in a given situation.
References


