Working Paper Proceedings

Engineering Project Organization Conference
Cle Elum, Washington, USA
June 28-30, 2016

Nudging for Smart Construction: Tackling Uncertainty by Changing Design Engineers' Choice Architecture

Marc van Buiten, University of Twente, the Netherlands
Andreas Hartmann, University of Twente, the Netherlands
Jeroen van der Meer, University of Twente, the Netherlands

Proceedings Editors
Jessica Kaminsky, University of Washington and Vedran Zerjav, University College London

NUDGING FOR SMART CONSTRUCTION: TACKLING UNCERTAINTY BY CHANGING DESIGN ENGINEERS' CHOICE ARCHITECTURE

Marc van Buiten[1], Andreas Hartmann[2], and Jeroen van der Meer[3]

ABSTRACT

Nudging refers to the deliberate change in choice architecture (i.e., the careful design of the environments in which people make choices) with the goal of engineering a particular (benign) outcome. The core proposition of nudging is that seemingly innocuous alterations of choice architecture matter: they influence decision making. Typically, the early stages of decision making in construction projects have a long-term impact on project performance, but are at the same time often fraught with uncertainties. We experimentally examine if tweaking a trade-off matrix used to assess preliminary design options in a design competition can raise uncertainty awareness of decision makers and can nudge decision makers away from riskier, more uncertain options, towards less risky and uncertain options. First results indicate a statistically unreliable shift in the predicted direction. Based on our research results, we suggest and discuss other pertinent and potentially more effective nudges.

KEYWORDS: nudging, decision making, design uncertainty, tender, contractors

INTRODUCTION

In recent years, contractors have increasingly been in charge of the design, building and maintenance of large infrastructure projects through integrated contracts. Using integrated contracts results in a design responsibility shift from the client to the contractor, i.e., the contractor becomes responsible for a larger part of the project life cycle. Although integrated contracts create opportunities for life-cycle oriented design optimizations, contractors need to anticipate the long-term effects of design decisions during the tender phase which places large information needs on tender teams. However, due to the competitive nature of the tender phase with its time, resource and budget constraints, tender teams are often forced to take design decisions without completely knowing the entire infrastructure requirements, its environment of operation, future design decisions, and emergent infrastructure behavior. Designers are often not aware of the uncertainties involved in their design decisions leading to design choices that may constrain decisions at later stages and compromise the project outcomes. Being able to identify uncertainties and incorporate them into the early stages of decision making is, on the one hand, beneficial to prevent any unexpected and undesired design effects, but is, on the other hand, also a challenge for many contractors.

[1] Researcher and lecturer, Department of Construction Engineering & Management, University of Twente, P.O. Box 217, 7500 AE Enschede, m.vanbuiten@utwente.nl.

[2] Associate Professor, Department of Construction Engineering & Management, University of Twente, P.O. Box 217, 7500 AE Enschede, a.hartmann@utwente.nl.

[3] PhD student, Department of Construction Engineering & Management, University of Twente, P.O. Box 217, 7500 AE Enschede, j.p.vandermeer@utwente.nl.

As part of an ongoing research project, we experimentally examine solutions to increase designers' awareness of uncertainties involved in design-decisions for large infrastructure projects. In particular, we focus on increasing designers' awareness of uncertainties at an early design stage, and minimizing human bias in reasoning about uncertainties, ultimately generating better design choices. Our approach to identify possible solutions is grounded in behavioral economics, a field that draws on many decades of research—most notably by Nobel laureate Daniel Kahneman and the late Amos Tversky—examining human errors and biases in reasoning about problems that involve risk and uncertainty. *Nudging,* in particular, has been highly influential in both academic and non-academic circles ever since economist Richard H. Thaler and legal scholar Cass R. Sunstein published the book *Nudge: Improving Decisions about Health, Wealth, and Happiness* (Thaler & Sunstein, 2008). It refers to the deliberate change in choice architecture (i.e., the careful design of the environments in which people make choices) with the goal of engineering a particular (benign) outcome.

In this paper, we report on a pilot experiment to examine choice outcomes under different representations of a decision aid (trade-off matrix) for selecting preliminary designs for further elaboration in a design competition that bears some similarity to the tender phase. The baseline condition encompassed a standard, albeit simple, trade-off matrix to assess preliminary designs. The alternative condition was created to nudge decision makers towards the safer option. Although results indicate a statistically unreliable shift in the predicted direction, the experiment provides a platform to discuss potentially more effective nudges.

BACKGROUND

Uncertainty in tendering large-scale construction projects

The realization of construction projects start with the definition of the customers' needs by the future owner of the construction. In most projects, a preliminary solution based on the customers' needs is publicly procured by the owner. Contractors have to develop a competitive bid for this given problem, for which there is no easy solution or obviously right answer. The best competitive bid can be found by iterating between many potential alternatives. Limited available time and resources, due to the procurement, hamper this iterative design process. In this iterative design process, engineers have to find the most appropriate solution by examining many potential alternatives using comprehensive technical knowledge and judgment. A structured and easy-to-use multi-criteria tool is used by many contractors to choose between alternatives. Criteria are typically based on requirements, cost and schedule.

A recent study about design challenges in construction tenders suggests that the design task is based on expert judgment and experience, and that the added value of potential alternatives is not always explicated when making design decisions (Van Der Meer, Hartmann, Van Der Horst, & Dewulf, 2015). The process of settling on the preferred solution ideally requires an assessment of alternatives at the level of the entire system. Unfortunately, clearly defined decision criteria at the system level are hard to come by. Furthermore, the competitive nature of tenders challenges engineers to make design decisions based on incomplete design specifications (Laryea, 2013). Engineers need to handle a dynamic construction environment and as well as capricious customer's needs in their quest for a satisfying design alternative. Each domain specific expert has to make difficult trade-offs that involve identifying and synthesizing alternatives to achieve desirable outcomes on sets of conflicting goals. During the decision-making process, individual engineers assimilate information differently before emphasizing particular factors and score design alternatives using multi-criteria analyses. Against this background, there is an apparent need for engineers to understand the nature of the uncertainties that are involved.

Creating choice architecture

Ideally, choice behavior is supposed to be invariant to various seemingly trivial factors, such as phrasing specifics (i.e., description invariance), contextual details (i.e., contextual invariance), and particulars of elicitation method (i.e., procedural invariance). But, of course, these invariances do not hold up that well in practice. People are often influenced by the words chosen to convey a message, as well as the details of the decision context and the way preferences are elicited. In short, our traditional normative depiction of decision making ("what we should do") is often at odds with a descriptive assessment ("what we actually do").

A natural question to ask is how to close this gap, i.e., how to render actual decision making more normatively defensible. This prescriptive approach ("what we can and should do") has been pursued by many researchers and professionals. Well-known is work on decision analysis by contributors such as Ron Howard, Ralph Keeney, Howard Raiffa, and John Hammond, although this is only the tip of the iceberg.

Sometimes a further distinction is made between *de*biasing and *re*biasing, depending on whether efforts to align descriptive behavior with normative standards focus on eliminating underlying cognitive biases altogether or merely counteract them (e.g., Larrick, 2004; Soman & Liu, 2010; also see Van Buiten & Hartmann, 2013). Nudging exemplifies the latter approach.

Nudges are not mandates and should be easy and cheap to avoid. The idea is to steer behavior, while retaining freedom of choice (sometimes dubbed libertarian paternalism). For example, displaying fruit at eye level to grab people's attention to promote a healthy diet is a nudge. Banning junk food altogether is not (Ly, Mazar, Zhao, & Soman, 2013). Other nudging examples include the use of default "opt-out" systems rather than "opt-in" systems to increase organ donation rates (e.g., Johnson & Goldstein, 2003), placement of green footprints on the pavement leading pedestrians to nearby garbage bins to reduce littering (e.g., Ly et al., 2013), and depictions of a fly in urinals to reduce "spillage" (e.g., Thaler & Sunstein, 2008). Recent reviews (Johnson et al., 2012; Ly et al., 2013) highlight these and other nudges that have been effectively used in recent years.

RESEARCH DESIGN AND METHOD

The core proposition of nudging proponents is that seemingly innocuous alterations of choice architecture matter: they influence decision making. We examine if tweaking a trade-off matrix used to assess preliminary design options in a design competition can nudge decision makers away from riskier, more uncertain options, towards less risky and uncertain options. Essentially, we examine whether we can capitalize on people's sensitivity to variations in elicitation method (i.e., violation of procedural invariance alluded to in the background section) to nudge decision makers towards more cautious behavior in tenders. To this end, an experiment was conducted that involved a hypothetical, but realistic, scenario. Care was taken to create a scenario that would both appeal to student participants (allowing them to create a vivid image of the situation), and be relevant in its implications for construction industry. The basic approach captured in this experiment is modeled after the approach of one of the larger Dutch construction companies. Of course, the version presented below is a highly stylized version of that approach (also see "Discussion and conclusion"). Our intention was not to change the original choice architecture, but to test in more general terms whether a nudge is able to raise risk awareness and will lead to a different choice outcome.

Participants

Sixty-three students from the bachelor program in civil engineering technology at University of Twente in Enschede participated in the in-class experiment. Thirty-two of them were students enrolled in a third-year mandatory methodology course that was part of their preparation work for their final thesis. They received no financial compensation for their participation. Their sole "compensation" consisted of receiving feedback on this experiment and the role of experimental methodology in civil engineering more generally in a subsequent session. Remaining students were first-year students enrolled in a course on traffic and transportation (selected for its reliably high attendance rate) who were willing to volunteer some time to participate in our experiment. Students from both sub-samples were randomly assigned to experimental conditions. The experiment took only a couple of minutes to complete.

Design and procedure

The experiment was conducted using paper and pencil[4]. Before distributing the experiment among participants, they were told that they were participating in a small study and that everything should be clear from the instruction page that was included. Participants were asked to be completely silent until everyone completed the study, and to provide an answer even when in doubt.

The instruction page asked participants to read the text carefully. They were told that they would be required to imagine (to the best of their ability) a hypothetical situation. It was stipulated that there was no time limit, and no wrong or right answer. If a question would arise, they could simply raise their hand. They were asked to turn off devices that might disrupt the study, and they could list their student number if they wished, but they could also choose to remain anonymous.

The two subsequent pages contained the scenario and a single question. The scenario concerned a design contest for students. Supposedly, recent increases in migration had the central agency responsible for "*the reception, supervision and departure (from the reception location) of asylum seekers coming to the Netherlands*" (COA, 2016) exploring innovative

[4] An online, or at least, digital format was considered. However, main advantages of control and quick data availability did not outweigh some serious disadvantages. For example, Google Forms—an obvious choice for collecting survey-type responses—has only limited layout and formatting options. Moreover, its use as a tool for experimentation requires finding a way to distribute different versions of a questionnaire among participants in real time (i.e., in the class room). And, of course, there is the additional drawback of technical failure risk (including coming across students that do not have a smartphone, laptop, or other means of completing an online questionnaire available).

solutions to the housing problem[5]. Participants were told that they were part of a student team that already had two rough ideas. The first option was to transform existing structures to be used as a home for asylum seekers. The other one would be to build something new from scratch. For display purposes these options were labeled option 1 and 2 respectively and two pictures were included (one depicting an abandoned office building, the other showing a plot of land). One of these ideas had to be developed further. In order to determine which, these students supposedly adopted a structured approach in the form of a trade-off matrix. This matrix accounted for six criteria COA explicated in the contest. These criteria (flexibility, time-to-market, sustainability, affordability, spatial and social quality, and risks and uncertainties) were briefly described. The presented matrix contained the scores of both options on the aforementioned criteria. A five-point scale was used ("- -", "-", "0", "+". "++"). Option 1 scored high on most criteria, but entailed substantial uncertainty. Option 2, although scoring a bit lower on most criteria, constituted the safer bet. Participants were asked to consider the matrix, and to indicate their preference on a binary scale ("Option 1", "Option 2").

Crucially, about half the participants received a version that included a standard, albeit very simple, trade-off matrix (condition 1—baseline), whereas the remaining participants received a "nudged" version (condition 2—nudge). One of the authors of this study handed out one version to students seated in the front rows of the class room. A research assistant distributed the other version among students seated in the back of the class room.

We employed several ideas from the literature on nudging to try to induce a preference reversal between conditions. Specifically, the idea was to nudge people away from the relatively attractive yet uncertain prospect (Option 1) to the more average but less uncertain prospect (Option 2). To this end, we reviewed (Johnson et al., 2012) for potentially effective nudges. We ran into a common criticism of the underlying literature, i.e., that "there are too many potential levers to change behavior, without a clear indication of their relative effectiveness" (DellaVigna & Pope, 2016). For our pilot experiment, we decided to focus on what Johnson et al. (2012) refer to as attribute parsimony and labeling. As a result, our "composite" nudge comprised four aspects. First, the risk and uncertainty criterion was moved to the first slot in both the matrix as the preceding description of the criteria. Second, a visual aid was used to increase prominence of the risk and uncertainty criterion. The "- -" score of Option 1 was complemented by a sad, red-colored emoticon. Similarly, the neutral "0" score of Option 2 was augmented with a yellow, neutral emoticon. Third, the risk and uncertainty criterion was unpacked by explicitly providing three examples. Lastly, the last snippet of text before the trade-off matrix and the title of the matrix contained additional descriptors referring to the choice involving risk and uncertainty. Figure 1 presents the two versions of the trade-off matrix as used in the two conditions.

[5] This scenario was based on an actual design contest from the central agency (see COA & Rijksbouwmeester, 2016).

Panel A: Baseline condition

Trade-off matrix

	Option	
Criterion	**OPTION 1 ("Conversion")**	**OPTION 2 ("New development")**
Flexibility	+	0
Time-to-market	+	+
Sustainability	++	+
Affordability	+	+
Spatial and social quality	++	+
Risks and uncertainties	--	0

Panel B: Nudge condition

Trade-off matrix for decision making under great uncertainty

	Option	
Criterion	**OPTION 1 ("Conversion")**	**OPTION 2 ("New development")**
Risks and uncertainties such as • Difficulties during site preparation • Failure to obtain permits • Objections residents	-- 	0 
Flexibility	+	0
Time-to-market	+	+
Sustainability	++	+
Affordability	+	+
Spatial and social quality	++	+

Figure 1. Trade-off matrices (translated from Dutch).

We pre-tested attractiveness of choice options as presented in the baseline condition. It was important to the study to have more or less equally attractive options in the baseline condition—or at least prevent an extremely attractive Option 2—as we needed scope the nudge overall preference from Option 1 towards Option 2. Three independent raters (all students pursuing a professional doctorate at our department) were asked to act as trial run participants. Upon completion, they were asked about (1) general inaccuracies and typos, and (2) their own preference and prediction of overall preferences in the student population. Their responses allowed to correct one typo and gave us no reason to change scores of the criteria for both options. It appeared that the preset scores made options approximately equally attractive; option 1 arguably even slightly more attractive than option 2.

In sum, the experiment had a one–way between–subjects design with two conditions (baseline vs. nudge). The single dependent variable was option choice (Option 1 vs. Option 2)[6].

RESULTS

Results of the experiment are displayed in Table 1. Across conditions, 47 out of 63 (75%) participants preferred the good but uncertain (option 1) over the more average but safer option (option 2). In the baseline condition 24 out of 30 (80%) chose option 1, whereas 23 out of 33 (70%) did so in the nudge condition. The slight preference shift towards the safer option was not statistically reliable as assessed by a z-test for proportions ($z = 0.94$, $p = 0.17$, one-tailed).

Table 1. Number of participants preferring option 1 or option 2 as a function of condition (baseline vs. nudge). Choice proportions are within parentheses.

Condition	Selected option		Total
	Option 1 (“Good but uncertain”)	Option 2 (“Average, but safe”)	
Baseline	24 (80%)	6 (20%)	*30*
Nudge	23 (70%)	10 (30%)	*33*
Total	*47 (75%)*	*16 (25%)*	*n=63*

Using standard statistical logic, these results tell us that the likelihood of finding the observed pattern of responses is not low enough (convention is to use $p = 0.05$) to reject the notion of an inconsequential nudge. We simply have insufficient reason to rule out the possibility that the nudge is ineffective, but we cannot conclude that it is ineffective either. The nudge might be ineffective, or we might simply be lacking power to rule out this possibility.

To assess whether the nudge is truly ineffective, confidence intervals can sometimes be informative (Aberson, 2002). A narrow confidence interval (CI) around the difference in proportions (choosing option 1 over option 2) between conditions would suggest—although not technically support—that the difference is of no real practical relevance. The 95% CI for the difference is (-0.11, 0.30)[7]. This is not a particularly narrow interval. In other words, the

[6] Experimental stimuli are available upon request (Dutch language only).

[7] The calculated 95% confidence interval for the choice proportion in the baseline condition is (0.63, 0.90). For the nudge conditions the interval is (0.53, 0.83).

difference in choice proportion could be quite substantial. The lack of precision in the interval prevents us from drawing meaningful conclusions.

DISCUSSION AND CONCLUSION

Results of this pilot experiment do not show a statistically reliable effect of our proposed nudge. It is, of course, conceivable to find an effect with a larger sample size (and increased power). Notwithstanding, nudging also has its limits and it would be useful to think about factors that might impede nudging, and ways to increase its effectiveness.

The strong overall preference for the good but uncertain option might weaken our result. Strong preferences will be less susceptible to nudging attempts than weak ones. Moreover, in hindsight, the placement of the smile close to the score ("- -", and "0") could have been a bit unfortunate. Instead of bolstering the score, the smile might have obscured it so that the risk/uncertainty criterion figured less, rather than more prominently. In addition, despite our efforts to increase salience of the risk/uncertainty criterion (e.g., by using a "composite" nudge), participants might still underestimate the effect of these risks materializing in practice. For example, failure to secure the required permits will adversely affect all other criteria. By presenting risk and uncertainty as seemingly independent criteria the relationship with other criteria might be obscured and the favorable evaluation of these criteria may still dominate the final choice. Future research could make this link more noticeable by adopting a different representation of risk in the trade-off matrix (e.g., by eliminating risk/uncertainty as a separate criterion and instead using intervals for scoring all the criteria). One could also stress the potential impacts of risks by including cautionary examples of other projects. Lastly, defaults have proven an effective nudging tool. Future research could include, e.g., references to choice behavior of peer groups that would thus serve as a default.

We supplemented our suggestions with responses obtained in the debriefing session we had with student participants. We asked participants about possible reasons for the lack of preference shift. Of particular interest were answers from nudge condition participants. To be sure, we are notoriously bad at appraising our own choice processes (e.g., West, Meserve, & Stanovich, 2012). More often than not, asking people to offer some insight into their decision making process will only trigger post-hoc rationalizations detached from reality. Nonetheless, we decided to reach out to student participants. As expected, it turned out to be a hard question. Some participants stated that the examples of risks and uncertainty listed in the trade-off matrix were simply deemed to insignificant to be any kind of factor, thus reducing its importance and taking away much of the appeal of the safer choice option. It would be interesting to see whether our student sample (and the student population at large) is more tolerant of risks than professionals are. This could mean that the nudge might be more effective with professionals than with students. More generally, population characteristics (e.g., technical ability, background knowledge) are of interest. In a future experimental setting the measurement of the uncertainty awareness and behavior may provide additional insights in nudge effectiveness and decision makers' attitude towards risk and uncertainty.

Future research would also have to manage to include more of the real-world complexity that has been intentionally excluded from our pilot study. Future work should try to better represent actual design processes. It is clear that designers do not make choices in isolation (i.e., individually), nor that they simply have to choose between a limited number of predefined designs. Moreover, path dependencies, learning, and status quo bias all influence the decision making dynamic.

All in all, it still seems to us that nudging can be of great value in construction management, but that the devil is in the details. The very nature of nudges poses limits on the severity of the alterations to the choice context that you can impose. Yet, these alterations still need to be strong enough to matter, i.e., to bring about a discernable difference in choice behavior.

Surprisingly, nudging has not yet received much attention in construction management (but see Shealy & Klotz, 2015; Van Buiten & Hartmann, 2013, for exceptions and related literature). Perhaps, in part, because nudges appear manipulative. There is something a bit aversive about tricking people into certain behaviors (e.g., marketing executives have long employed very similar practices to good financial effect). But as Johnson et al. (2012, p. 488) remind us:

"While it is tempting to think that choices can be presented in a "neutral" way ("Just the facts, Ma'am"), the reality is that there is no neutral architecture—any way a choice is presented will influence how the decision-maker chooses."

It would be ill-advised to pass up an opportunity to improve practices in the construction industry just because it sounds iffy. What the construction industry needs now are choice architects!

ACKNOWLEDGMENT

The authors would like to thank Samy Awad for his assistance conducting the experiment.

REFERENCES

Aberson, C. (2002). Interpreting null results: Improving presentation and conclusions with confidence intervals. *JASNH, 1*(3), 36-42.

COA. (2016). Web page Centraal Orgaan opvang Asielzoekers. Retrieved from https://www.coa.nl/en/

COA, & Rijksbouwmeester. (2016). A home away from home. Retrieved from http://www.nederlandwordtanders.nl/openoproep/open_oproep/

DellaVigna, S., & Pope, D. (Producer). (2016). Inducing effort with behavioural intervention *VoxEU.org – CEPR's policy portal*. Retrieved from http://voxeu.org/article/motivations-behind-effort-evidence-and-expert-forecasts

Johnson, E. J., & Goldstein, D. G. (2003). Do defaults save lives? *Science, 302*(5649), 1338-1339.

Johnson, E. J., Shu, S. B., Dellaert, B. G. C., Fox, C., Goldstein, D. G., Haubl, G., . . . Weber, E. U. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters, 23*(2), 487-504.

Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (pp. 316-338). Malden, USA: Blackwell Publishing Ltd.

Laryea, S. (2013). Nature of tender review meetings. *Journal of Construction Engineering and Management, 139*(8), 927-940.

Ly, K., Mazar, N., Zhao, M., & Soman, D. (2013, March 15). *A Practitioner's Guide to Nudging*. (Rotman School of Management Working Paper No. 2609347). Toronto: University of Toronto.

Shealy, T., & Klotz, L. (2015). Well-Endowed Rating Systems: How Modified Defaults Can Lead to More Sustainable Performance. *Journal of Construction Engineering and Management, 141*(10).

Soman, D., & Liu, M. W. (2010). Debiasing or rebiasing? Moderating the illusion of delayed incentives. *Journal of Economic Psychology, 32*(3), 307-316.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge : improving decisions about health, wealth, and happiness*. New Haven, Conn.: Yale University Press.

Van Buiten, M., & Hartmann, A. (2013). *Public-private partnerships: cognitive biases in the field.* Paper presented at the Engineering Project Organization Conf. 2013, Winterpark, CO.

Van Der Meer, J., Hartmann, A., Van Der Horst, A., & Dewulf, G. (2015). Challenges of using systems engineering for design decisions in large infrastructure tenders. *Engineering Project Organization Journal, 5*(4), 133-145.

West, R. F., Meserve, R. J., & Stanovich, K. E. (2012). Cognitive sophistication does not attenuate the bias blind spot. *Journal of Personality and Social Psychology*, *103*(3), 506-519.