

Draft: 7/15/17

**Choice Architecture 2.0:
Behavioral Policy as a Social Interaction**

Job M. T. Krijnen

David Tannenbaum

Craig R. Fox

Author Note

Job M. T. Krijnen, UCLA Anderson School of Management. David Tannenbaum, Management Department, Eccles School of Business, University of Utah. Craig R. Fox, UCLA Anderson School of Management and Department of Psychology. An earlier version of ideas contained in this paper was presented by Fox at the August 2016 Academy of Management Conference. We thank Hengchen Dai, Jana Gallus, and Stephen Spiller for helpful comments on an earlier draft of this paper. Correspondence concerning this article should be addressed to Job Krijnen: job.krijnen@anderson.ucla.edu.

Abstract

In this essay we propose a new framework, *choice architecture 2.0*, which construes behavioral policy interventions as an implicit social interaction between choice architect and decision maker. We argue that choice architects should be treated as endogenous to any behavioral policy intervention because decision makers engage in *social sensemaking*. Thus, choice architecture 2.0 adds an analysis of inferences that decision makers may draw about: (1) the beliefs and intentions of the choice architect; (2) the social meaning of their own behavior. We present examples to motivate choice architecture 2.0 as a framework for anticipating when policy interventions are likely to succeed or fail, and for identifying new policy tools. Additionally, we identify cues that may trigger social sensemaking by decision makers.

Choice Architecture 2.0: Behavioral Policy as a Social Interaction

In the fall of 2016, the government of the Netherlands wished to increase consent for organ donation by its residents. A chamber of the Dutch legislature passed a highly publicized bill that recommended changing the procedure from requiring explicit consent on an online registry (i.e., an *opt-in default*) to presumed consent unless residents explicitly decline (i.e., an *opt-out default*).^{1,2} This proposal was motivated by recent successes of policymakers applying insights from behavioral science to design choice environments that promote better decisions of targeted individuals — an activity that has come to be known as *choice architecture*.^{3,4} The strategic use of defaults has proven to be one of the most successful tools of choice architecture. For instance, employees are more likely to save for retirement when the default is automatic enrollment into a 401(k) plan,⁵ doctors are more likely to prescribe generic drugs when the default setting in an electronic health record system is to replace branded prescriptions with generic alternatives,⁶ and consumers are more likely to receive energy from renewable sources in their homes when they are automatically enrolled into a green energy plan rather than a grey energy plan.⁷ Of particular relevance, rates of consent for organ donation are much higher in European countries where consent is presumed compared to countries where residents must opt-in.⁸

Based on these findings, it seemed natural for the Dutch legislature to assume that the shift from explicit consent to presumed consent would succeed in increasing organ donation. Unfortunately, the month the bill passed the number of residents who explicitly registered as non-donors leaped to roughly 40 times the number for a typical previous month.⁹ This dramatic increase in active rejections was observed not only among newly registering residents, but also

among those who had previously consented to donation and then went to the trouble of going back online to revoke their consent.

Interestingly, the backlash observed by the Dutch to changing the organ donation regime is not an isolated incident. In the early 1990s, the states of Virginia and Texas saw rates of non-consent rise markedly after they switched their organ donation policies from explicit consent to mandated choice (where residents must register organ donation preferences when applying for or renewing a driver's license).^{10,11} Why did changing defaults backfire so spectacularly?

One possibility, which we explore in this essay, is that residents may have construed the (proposed) change in choice architecture as an attempt at coercion by their government. Residents may have recognized that policymakers altered the policy regime with the intention of increasing consent for organ donation, which provoked many people to rebuke this attempt by opting out.

If true, this interpretation suggests that we need to update our understanding of choice architecture. Policymakers would not only need to consider how features of a choice environment may affect the behavior of decision makers, but also how decision makers construe the beliefs and intentions of the policymakers themselves. Furthermore, policymakers may also need to consider how decision makers think their actions will be perceived by the choice architect and other observers. For instance, Dutch registered donors revoking their consent may have wanted to communicate a "protest" against the government's policy change. In short, these anecdotes suggest a need to reconstrue choice architecture as an implicit two-way interaction between the choice architect and decision makers.

From Choice Architecture 1.0 to 2.0

The traditional approach to choice architecture treats the choice architect as an exogenous agent whose role is to design the appropriate choice environment. Although the effectiveness of particular tools of choice architecture, such as defaults, are assumed to vary from context to context, the beliefs and intentions of the choice architect are treated as irrelevant. In this essay, we propose an update to the traditional framework in which targets of choice architecture are *social sensemakers* who actively seek to understand why the choice architect has presented them with a particular choice environment, and how their own actions will be evaluated by the choice architect and others. We call this approach *choice architecture 2.0*. This perspective may help choice architects better understand and predict contextual factors and implementation details that can critically influence the success or failure of an intervention. Choice architecture 2.0 may also suggest potent new tools for influencing behavior that are opaque when viewed through the lens of traditional choice architecture.

We begin with the working assumption that decision makers are often uncertain about their decisions (e.g., whether or not to consent to organ donation, which health care plan is best, how much energy to consume). To help make sense of the situation, individuals may glean information from the implicit conversation between themselves and the choice architect.^{12,13} They may try to discern what the choice architect is tacitly communicating (or revealing) through his or her selection of a particular choice architecture, and also consider what their own choices signal back to the architect.

In the remainder of this essay, we provide evidence for the insufficiency of choice architecture 1.0 and motivate the need for the 2.0 update. We first present several examples of how social sensemaking about the choice architect's beliefs and intentions and the decision

maker's own behavior can help or hinder the effectiveness of choice architecture tools. Next, we identify several factors that may trigger the social sensemaking process. Finally, we conclude with comments on how policymakers can incorporate a choice architecture 2.0 perspective into the design of policy.

Inferences About the Beliefs and Intentions of the Choice Architect

In the US, 401(k) plans have become a popular investment vehicle to help employees save for retirement. Nonetheless, many eligible employees fail to take advantage of the attractive opportunities that these plans may provide.^{14,15} Thaler and Benartzi¹⁶ found that a minor change in the setup of a 401(k) plan — providing people with the option to pre-commit to initiating and then escalating retirement contributions in the future, when receiving raises — boosted both participation and downstream saving. Surprisingly, a recent field study by Beshears, Dai, Milkman, and Benartzi¹⁷ found that offering employees the option to pre-commit to future saving led to a subsequent *decrease* in overall savings contribution rates. Why would seemingly identical interventions increase retirement saving in one case and decrease it in another?

As it turns out, there was a subtle but critical difference in how the intervention was implemented by the two research teams. Whereas Thaler and Benartzi first asked employees whether they wanted to start saving immediately and then presented the “save more tomorrow” option only to those who had not enrolled in the conventional plan, Beshears and colleagues provided employees with a direct choice between initiating (or increasing) saving today versus initiating (or escalating) saving later. Recognizing this procedural difference, Beshears et al. tested an explanation of why their version of the save more tomorrow program failed: presenting employees with both options side by side (rather than providing the option to delay only after

participants decided not to save immediately) was construed by employees as a signal from the employer that retirement saving is not particularly urgent. Many employees may reasonably assume that their employer knows more than they do about the urgency of saving for retirement, and that the choice environment was constructed with these considerations in mind. In this way, the choice architecture “leaked” information from the choice architect to the decision maker.^{18,19}

Information leakage can take many forms, can be intentional or unintentional, and can help or hinder the goals of the choice architect. Below we address five commonplace choice architecture tools in which decision makers appear to draw inferences about the beliefs and intentions of the choice architect from features of the choice environment: defaults, anchors, item position and order, menu partitions, and incentives.

Defaults. As discussed in the introduction, choice architects often designate an option as the default consequence when no action is taken by the decision maker. The strategic assignment of defaults is a powerful policy tool for several reasons,²⁰ one of which is that defaults are sometimes interpreted as an implicit endorsement by the choice architect, or as a signal of modal preference.^{21,22} It is well known that participation in 401(k) retirement saving plans increases when employees are automatically enrolled,^{23,24} and in one field study approximately one-third of employees who stayed with the default retirement plan indicated that they did so because they believed it to be the recommended option.²⁵ Both laboratory and field studies find that the more a default option is viewed as an implicit recommendation, the more likely people are to stick with that option.^{26,27,28}

When a decision maker construes defaults as recommendations, the competence and benevolence of the choice architect also become relevant concerns. A decision maker who distrusts the choice architect will tend to be skeptical of the options that the choice architect

appears to endorse. For instance, a consumer may reject an expensive upgrade package into which she is automatically enrolled if she feels the choice architect is endorsing the package merely to get the consumer to spend more money. Indeed, in both laboratory and field settings researchers have documented examples in which defaults selected by distrusted choice architects have backfired.^{29,30,31,32}

Anchors. When people make numerical evaluations they tend to be biased toward focal values or “anchors.” For instance, real estate agents are unduly influenced by the listing price when assessing the fair market value of a home.³³ Another example is the minimum-repayment information provided by credit card companies that was mandated in the US by the CARD Act of 2009. Whereas these disclosures were intended to help people avoid rising debt through compounding interest, various studies suggest that minimum-repayment information can actually cause people to make *lower* monthly repayments than they otherwise would because such minimum-repayment amounts act as anchors.^{34,35,36}

There are various psychological factors that contribute to anchoring effects.³⁷ In the context of behavioral policy, one factor that may enhance the strength of an anchor is that decision makers may view the focal number as a signal from the choice architect concerning what range of behavior would be sensible, appropriate, or recommended. One implication of this account is that anchoring effects will be stronger if the choice architect is perceived to be more benevolent and/or more competent. For example, in one set of studies³⁸ when real estate negotiators responded to opening offers (i.e., anchors) from sellers, their counteroffers increased with the perceived competence of the seller, which they inferred from the precision of sellers’ opening offers (e.g., a more precise opening offer of \$799,800 might lead a buyer to make a higher counteroffer than a less precise opening offer of \$800,000).

It is worth noting that social sensemaking underlying anchoring effects can sometimes undermine the intended effect of other tools of choice architecture. For instance, one set of studies on charitable giving³⁹ found that setting a low donation amount as the default led donors to give less money on average compared to if there was no default or if the default was a high donation amount. Interestingly, the downward pull of a low default amount was stronger when the default was presented as a “suggested amount” than when it was presented as a randomly generated default. Apparently, focal values presented as part of choice architecture may cause social sensemakers to infer that they are endorsements, which can sometimes cause them to suppress desirable behavior.

To illustrate this point further, consider the domain of retirement saving, where automatic enrollment into a low saving rate and conservative portfolio may cause employees who would otherwise save more money and save more aggressively to scale back their savings.⁴⁰ Because of this, such employees may accumulate less money at retirement compared to if they had been left to their own devices to opt-in to a particular rate and investment strategy. One way to avoid this pitfall could be to explicitly communicate to decision makers that the focal value (e.g., default saving rate, minimum-repayment amount) is not intended to be the recommended value.

Item Position and Order. Choice architects must decide how to position and order the options they wish to present to decision makers. For instance, a drop-down menu on a website may list forms of available health coverage; an election ballot lists candidates in a particular order. Studies have found that the positioning of options impacts what people choose, but the nature of such position effects vary across contexts. For instance, in supermarkets, consumers are biased to choose products in the center of a display;⁴¹ on food menus, people tend to favor options closer to the top or bottom;⁴² in elections, political candidates tend to receive more votes

when they are on the top of the ballot than when they are in any other position.⁴³ Bar-Hillel⁴⁴ suggested that these diverging findings can be explained by inferences that decision makers draw about the choice architect. Consumers may believe that a supermarket places its most popular product on center stage, whereas voters may think that the top-spot on the ballot is reserved for the incumbent or the most popular candidate. Consistent with this account, product position makes less of a difference when consumers believe that positioning is determined at random, thereby essentially blocking meaningful inferences about the beliefs and intentions of the choice architect.⁴⁵

Menu Partitions. Choice architects often must partition the list of available options into subsets or groups. For instance, those constructing menus for 401(k) plans may group available investments by geography (domestic versus international funds), by size (small cap versus large cap funds), or by risk profile (conservative versus aggressive funds). The manner in which the menu space is partitioned can have a profound effect on how people choose, even when the set of available options remains constant.⁴⁶ For instance, in one study participants recommended allocating far more financial aid to students from low-income families when they allocated a fixed budget over families in several lower-income ranges and one higher-income range {\$15,000 per year or less; \$15,001-\$30,000; \$30,001-\$45,000; \$45,001-\$60,000; \$60,001-\$75,000; more than \$75,000} compared to if they had been asked to allocate over one lower-income range and several higher-income ranges {\$75,000 per year or less; \$75,001-\$85,000; \$85,001-\$100,000; \$100,001-\$120,000; \$120,001-\$145,000; more than \$145,000}. In another study,⁴⁷ health care providers were presented with clinical vignettes and a list of possible medications, some of which were grouped (e.g. various antibiotic medications) and some of

which were listed separately (e.g., various non-antibiotic medications). Providers were biased to favor prescribing medications that were listed separately over those that were grouped together.

Although partition dependence has been amply observed in the absence of information leakage,^{48,49,50} recent studies have found that social inferences can sometimes play a substantial role in driving this phenomenon.⁵¹ In the absence of explicit information about the rationale for a grouping scheme, decision makers may infer that the choice architect assigned options to groups to reflect a fairly even distribution of the general population's preferences over these groups. This, in turn, can cause uncertain decision makers to gravitate toward choosing (ungrouped) options that appear to be more popular.⁵²

Incentives. Policymakers often introduce financial rewards or penalties as a way to promote desired behavior or discourage undesired behavior. In addition to the direct material inducement associated with rewards and punishments, the way in which financial incentives are presented can leak information about what the choice architect views as acceptable or appropriate conduct. For instance, incentives framed as punishments may signal a stronger moral condemnation of unwanted behavior than incentives framed as rewards, even when both framings lead to the same consequences.⁵³ In one study⁵⁴ participants learned about a company worksite wellness policy that introduced either a premium surcharge for its overweight employees or a premium discount for its healthy-weight employees. Even though the two policies were formally equivalent, participants inferred that the company introducing a surcharge (i.e., punishment) held especially negative attitudes about its overweight employees, whereas introducing a discount for healthy-weight employees (i.e., reward) was not viewed as signaling negative company attitudes. In a follow-up study, participants with higher Body Mass Indices reported that they would feel more stigmatized at work, and that they would be more willing to

consider looking for employment elsewhere, if their employer implemented an overweight penalty but not if their employer implemented a healthy-weight reward, precisely because only the overweight penalty communicated a negative message to these employees. Furthermore, a recent field experiment of students found greater willingness to vote against a penalty-based grading policy compared to a formally equivalent reward-based grading policy, in part because the penalty-based policy was perceived to communicate negative attitudes from the instructor.⁵⁵

A decision maker may also interpret the introduction of a reward or punishment as information from the choice architect about the attractiveness or difficulty of the targeted behavior.⁵⁶ Thus, a child who is rewarded by her parents for practicing the piano every day may infer that playing the piano is particularly uninteresting or difficult. In the policy domain, one study found that residents of two Swiss communities were less likely to accept the building of a nuclear waste facility near their homes when they were offered financial compensation should the facility be built in their area.⁵⁷ Presumably, residents inferred from the offer of financial compensation that having a nuclear waste facility nearby was especially hazardous or unattractive.

Inferences About the Social Meaning of the Decision Maker's Own Behavior

Overprescription of antibiotics by clinicians is a global health concern, as it contributes to the spread of antibiotic-resistant “superbugs.”^{58,59} In an attempt to reduce inappropriate prescription of antibiotics by U.S. clinicians, Meeker and her colleagues⁶⁰ examined the effectiveness of three behavioral interventions, two of which dramatically curtailed inappropriate prescribing for non-bacterial upper respiratory infections: compared to a control condition, clinicians became much less likely to prescribe antibiotics if they were prompted by the

electronic health record system to provide a written justification for each dubious prescription, or if they were provided monthly e-mails comparing their inappropriate prescription rates with “top providers” in their region. However, there is a remarkable side-note to this success story. Clinicians in the control group — who received no intervention beyond a bland education module — *also* reduced their inappropriate antibiotic prescribing substantially over the course of the study, by nearly one-half. It seems unlikely that the education module was responsible for this reduction, as it presented little that was new and previous educational interventions had not generally been effective in this context.⁶¹ Also, during the intervention period there were no special independent attempts by practices in the control condition to reduce antibiotic prescribing, so this does not appear to be the cause. Why would the control condition have such a strong impact on prescribing behavior?

Choice architecture 2.0 refocuses attention on what might otherwise be seen as an experimental artefact under a traditional choice architecture 1.0 framework. If we treat clinicians as social sensemakers, it seems plausible that most clinicians in the trial attempted to infer the goals and expectations of the researchers, along with what would be considered desirable behavior in this setting. Knowing that their prescribing behavior would be monitored, clinicians may have adjusted their behavior in order to be seen in the best possible light by the choice architect (who in this case were researchers from several prestigious institutions, including Harvard Medical School). Indeed, it is worth noting that the most precipitous reduction in antibiotic-inappropriate prescribing occurred at the very beginning of the intervention, and that these effects persisted throughout the intervention period.⁶²

Below we examine how inferences about the perceived social meaning of the decision-maker’s own behavior can influence the effectiveness of choice architecture interventions, using

examples from research on accountability, defaults, and incentives. Naturally, this secondary stage of social sensemaking will generally follow a primary stage in which the beliefs and intentions of the choice architect are considered, as described earlier.

Hawthorne and Accountability Effects. The strong impact of Meeker and colleagues' control intervention is not an isolated incident. Field studies often find participants behaving more desirably when they know they are being monitored. For instance, one study found that airline captains improved their fuel efficiency decisions even in the control condition in which they merely learned that researchers were observing them.⁶³ Another study found that households reduced electricity usage after receiving a postcard informing them that they had been selected as participants in a study on electricity usage.⁶⁴ These so-called "Hawthorne effects" are typically viewed as an empirical nuisance that challenges researchers' ability to assess the independent effects of experimental interventions.^{65,66} However, from a choice architecture 2.0 perspective, Hawthorne-like effects are not merely experimental artefacts; they represent potentially powerful and cost-effective tools of choice architecture.

Hawthorne-like effects may be a particularly subtle form of a family of interventions. Decision makers' concern with the social meaning of their actions, and resulting shifts toward socially desirable behavior, can be amplified when a choice environment makes them feel particularly accountable to observers. For instance, one study found that promising to publish neighborhood voting records (i.e., who voted and who did not) increased turnout in Michigan's 2006 primary election,⁶⁷ and a vast literature suggests that accountability or reputational concerns may be more salient when decision makers expect that the choice architect will ask them to provide justification for their choices.⁶⁸

We hasten to add that although decision makers may try to act in socially desirable ways to please a respected or trusted choice architect or other observer, in some contexts they may instead react *against* the perceived intrusion of a choice architect. A long tradition of research has established that in order to maintain and protect control over their own lives, people may actively resist persuasion attempts⁶⁹ or act against threats to their freedom to choose.^{70,71,72} In some cases decision makers may perceive choice architecture as coercive or autonomy-threatening, which may motivate them to demonstrate their independence by behaving in opposition to what they surmise to be the choice architect's goal — as we discussed in the cases of changes in organ consent default regimes in the Netherlands, Virginia, and Texas. Two further notes are in order in connection with reactance against choice architecture. First, although it has been asserted that reactance against choice architecture is unlikely when autonomy is preserved,⁷³ we argue that such effects can and do occur on occasion, especially in domains that are personally consequential,⁷⁴ with people who are especially concerned about their sense of freedom,⁷⁵ and in situations where the choice architect is distrusted.^{76,77} Second, we note that reactance against choice architecture need not be driven only by concerns about signals that are broadcast to others. For instance, in the aforementioned cases of reactance to changes in organ donation consent defaults, the “protest” was generally anonymous and the primary motivation of many protestors may have been to signal to *themselves* that they are autonomous.

Defaults. Changes in a default regime — for instance, switching from an opt-in to an opt-out system for organ donation — may not only communicate information from the choice architect to the decision maker as mentioned earlier, but may also affect the perceived social meaning of choices made by the decision maker. For instance, Germany has much lower consent rates for organ donation than Austria, a difference that is commonly attributed to the fact that

Germany uses an explicit consent (opt-in) default whereas Austria uses a presumed consent (opt-out) default.⁷⁸ Not coincidentally, Germans also assign more meaning to the act of organ donation than do Austrians, despite strong cultural similarities between the two countries. Likewise, American observers rated the act of organ donation in opt-in countries as comparable to self-sacrificing acts such as bequeathing one's wealth to charity, whereas they rated the act of organ donation in opt-out countries as less significant, comparable to courteous behaviors such as letting another person go ahead in line.⁷⁹

Such presentational concerns can affect the success of default interventions, especially when choices could reveal sensitive or stigmatizing information about the decision maker. In one study,⁸⁰ participants had the opportunity to test for a (fictitious) disease that they had read about during a previous testing session on health characteristics. Some participants learned that the disease was contracted due to unprotected sex (a socially stigmatized behavior) whereas others learned that the disease was transmitted through commonplace behaviors such as coughing or sneezing. Furthermore, the option to test for the disease was either presented to participants in the form of an opt-in default (where testing is voluntary) or as an opt-out default (where testing is routine but one can opt-out). Interestingly, the default intervention was especially powerful when testing involved a disease with potentially stigmatizing implications. Participants were more reluctant to opt-in to voluntary testing for a stigmatizing disease because of what their choice might reveal ("getting tested may imply that I've engaged in such behaviors and have something to worry about"). For similar reasons participants were also reluctant to opt-out of routine testing ("not getting tested may look like I have something to hide"). Thus, the social meaning that decision makers anticipate will be ascribed to their acts of opting in or opting out can have a substantial influence the effectiveness of default interventions.

Incentives. As discussed earlier, people often infer information about the choice architect from the framing of incentives as a reward or punishment. What's more, the selected incentive structure can also affect the decision makers' inferred social meaning of their own behavior.⁸¹ As a result, well-meaning incentives can sometimes backfire.^{82,83} Consider a well-known field study run at a daycare center,⁸⁴ which found that the introduction of a small fine for parents who were late to pick up their children led to an *increase* in the rate of tardy parents. Researchers speculated that introducing the fine changed the perceived meaning of showing up late from a moral violation (or perhaps proof of bad parenting) to a mere financial transaction. Apparently, the small fine was construed by most parents as a price rather than a penalty so that it ironically licensed them to freely take advantage of additional childcare.

A similar "crowding-out effect" can occur when a choice architect attempts to boost prosocial behavior by introducing a reward. For instance, in one study people were less likely to donate money to charities when greater attention was directed toward a gift they would receive from the charity if they donated.⁸⁵ People may infer from the introduction of a reward that their behavior signals self-interest instead of altruistic motivations to observers. Consistent with this explanation, rewards have been found to increase prosocial behavior when the outward appearance of self-interest is weakened, as when a reward for blood donation takes the form of a voucher,⁸⁶ a non-monetary gift,⁸⁷ a paid day off work,⁸⁸ a symbolic public award (e.g., a medal),⁸⁹ or when people can donate the money they receive to charity.⁹⁰

Triggers of Social Sensemaking

We have argued in this essay that a more complete account of choice architecture must incorporate the explicit or implicit presence of a choice architect, because decision makers may

attempt to infer the choice architect's beliefs and intentions from the particular selection of a choice environment as well as what the decision maker's own behavior might communicate back to the choice architect or other observers. While it has been asserted that there is no neutral way to present a decision and that choice architecture plays a role in all decisions,⁹¹ we do not expect that decision makers will engage in a social sensemaking exercise in every choice environment they encounter. For one thing, if social sensemaking requires conscious and deliberate thinking then such effects should only arise when decision makers have sufficient time and cognitive resources available. For instance, a rushed diner engaged in a business conversation may be too distracted to explicitly consider the restaurant's rationale for the particular grouping and ordering of menu items. This said, there may also be instances where social sensemaking is intuitive, fast, and automatic. For instance, if people learn from repeated experience that popular items are typically placed near the top of a restaurant's menu, they may come to automatically associate primacy with popularity, even when they lack cognitive resources for effortful deliberation. Future research is needed to determine the extent to which social sensemaking is conscious or automatic in different choice architecture contexts. For now, we assert that there are several factors, described below, that tend to trigger decision makers to engage in social sensemaking when they have sufficient cognitive resources. Understanding these triggers can assist choice architects in designing effective interventions.

Importance. We expect that decision makers are more likely to consider the goals and motives of a choice architect when the decision is personally relevant or important, as with commonplace social attributions.⁹² Moreover, dual-process theories of persuasion posit that people are more likely to engage in careful elaboration (or more systematic processing) of messages from others when they are motivated by high personal relevance or importance.^{93,94} For

these reasons, we assert that decision makers are more likely to derive social inferences from choice architecture associated with particularly consequential decisions. For instance, it seems plausible that employees are more likely to think about their employer's reasons for selecting a default retirement saving program than the reasons behind the default font size used by their employer's email program.

Abnormality or Change. People are more likely to draw inferences when they notice something unusual, unexpected, or statistically uncommon,^{95,96} which requires a comparison relative to a norm, an alternative, or a counterfactual.^{97,98} We therefore expect that decision makers are more likely to engage in social sensemaking about choice architecture when they recognize that the choice could have been presented in a different way. This may be especially apparent when a key aspect of the choice environment has recently been changed or when a new policy is introduced. For instance, most Dutch residents probably noticed the highly publicized proposal by their government to change the organ donation default policy from explicit consent to presumed consent, setting in motion attempts by many residents to infer the government's intention behind the policy change. Such inferences are less likely among residents of countries where the choice architecture represents a continuation of the status quo (e.g., neighboring Belgium, where a presumed consent default for organ donation has been the norm for many years). We note that in cases where sensemaking is triggered by a change in choice environment, people may gradually come to regard the new policy as normal and therefore be less likely to engage in sensemaking over time.

Preference uncertainty. A decision maker who feels relatively uncertain about the decision at hand may actively search for cues in the choice environment to guide his or her choice. This may partly explain why defaults in retirement saving have a greater impact on

people with little financial knowledge.⁹⁹ Although it may be rational for decision makers to look for guidance when they are uncertain about the proper course of action, we emphasize that subjective knowledge — the feeling of knowing — need not coincide with the decision maker’s actual (objective) level of knowledge. For instance, people feel less knowledgeable when they are made aware that they are lacking pertinent information or when they are reminded of others with greater expertise.^{100,101} Indeed, a recent series of vignette studies found that people who were made to *feel* subjectively less knowledgeable (without affecting their actual level of knowledge) were more receptive to adopting the default retirement saving option.¹⁰² We hasten to add that the tendency for preference uncertainty to increase the impact of choice architecture on choices may operate not only by increased triggering of social sensemaking but also through increased reliance on cues discerned through social sensemaking; further research on this topic is needed.

Distrust or Skepticism. As alluded to earlier, distrust of the choice architect may trigger a decision maker to more carefully consider the choice architect’s goals and motives. This may, in turn, reduce the effectiveness of policy interventions. Studies find that decision makers are more disapproving of behavioral interventions when these are implemented by choice architects whom they oppose politically,¹⁰³ or whom they perceive as dishonest.¹⁰⁴ This may explain why in one field experiment households in more politically conservative counties were more resistant to “green nudges” such as the use of feedback and social norms to promote energy conservation.¹⁰⁵ We suspect that regional differences in trust in government could be one factor explaining variation in the level of public acceptance of behavioral policy interventions.¹⁰⁶

We also note that past experiences with choice environments associated with a particular choice architect may affect trust in the architect, and that this may in turn influence the impact of

associated choice architecture in the future. For instance, a choice architect who sets a “smart default” tailored to a decision maker’s preferences is likely to improve the decision maker’s outcomes,¹⁰⁷ which may increase the decision maker’s trust in the choice architect and thereby increase his or her willingness to rely on future defaults from the same choice architect. Indeed, a recent study finds that participants who were presented with a high-quality default option were more likely to rely on the default in subsequent rounds of the same decision task than participants who were presented with a randomly selected default option.¹⁰⁸

Transparency. People are more likely to engage in social sensemaking when the (strategic use of) choice architecture is explicitly pointed out to them. While this may seem obvious it is both important and timely, as new tools of choice architecture have recently been debated by scholars, policymakers, and journalists. Many have argued that the strategic use of choice architecture is more ethical and more acceptable to the public when it is made more transparent to them.^{109,110} A choice architecture 2.0 lens not only reveals that transparency will tend to trigger social sensemaking, but also helps identify the likely magnitude and direction of its impact on the success of choice architecture. For instance, some research has found that even when the choice architects (i.e., the researchers) were transparent about the typical impact that defaults have on people, these effects did not diminish, presumably because default effects in these studies were driven by cognitive factors rather than social inferences (endorsements do not appear to be particularly relevant in those contexts).^{111,112} This said, studies on resistance to persuasion suggest that when transparency highlights the choice architect’s *intention* to nudge behavior, decision makers may shift their attitudes (and therefore choices) in the opposite direction.^{113,114} On the other hand, honesty about the goals and motives underlying choice architecture has also been found to reduce skepticism, increase perceived fairness, and engender

trust.^{115,116} It therefore seems plausible that in the long run transparency can boost effectiveness of many instances of choice architecture, but further research is needed on this topic.

Concluding Remarks

The impact of choice architecture on decisions is not always easy to anticipate. In this essay, we proposed an update — from choice architecture 1.0 to choice architecture 2.0 — that we believe can help policymakers better understand and predict in what contexts a particular intervention may succeed or fail. Choice architecture 2.0 enhances the traditional framework by treating the choice architect as endogenous so that the decision maker sometimes seeks information about the beliefs and intentions of the choice architect, and considers the information that his or her behavior may communicate to the choice architect and other observers. This updated perspective can help policymakers design and implement more effective choice architecture by highlighting seemingly irrelevant implementation details and contextual factors that critically influence the success of an intervention, and by suggesting promising new tools of choice architecture.

Returning to the organ donation anecdote with which we began this essay, one may ask how the choice architecture 2.0 framework could have informed the Dutch legislature and prevented the backlash. First, choice architecture 2.0 would have made it clear that a change in choice architecture will prompt many residents to consider the intention of the Dutch legislature (which was to increase consent to organ donation). Had the Dutch government anticipated that many residents would react against a perceived attempt to manipulate them into becoming organ donors, government officials might have taken steps to preempt such backlash by more carefully managing their communication about the proposed policy change. For instance, rather than

speaking about the goal of nudging more residents to become organ donors by changing the default to presumed consent, legislators might have emphasized that because most Dutch residents wish to become potential organ donors, the bill was designed to reduce obstacles to residents taking such action. Such an approach might have been more successful for two reasons. First, it signals a descriptive social norm (most Dutch residents prefer to be potential donors) toward which people may gravitate. Second, it frames the policy change as one designed to help residents express their preferences rather than nudge them to do something the choice architect deems desirable, and such rationales are likely to reduce reactance. Of course this analysis is meant to be illustrative and would need to be confirmed empirically. More generally, we hope that the choice architecture 2.0 perspective advanced in this essay will inspire a fruitful stream of future research that traces the relationship between social sensemaking by decision makers and the effectiveness of behavioral policy interventions.

¹ Back, M. (2016, September 15). 4.495 mensen trekken donorregistratie in [4.495 retract donor registration]. *NRC*. Retrieved from <http://www.nrc.nl/>.

² Kuiken, A. (2016, September 15). 18.890 Nederlanders zeggen ‘nee’ tegen Donorregister [18.890 Dutch say ‘no’ to donor registration], *Trouw*. Retrieved from <http://www.trouw.nl/>.

³ Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about wealth, health, and happiness*. New Haven, CT: Yale University Press.

⁴ Johnson, E. J., Shu, S., Dellaert, B., Fox, C., Goldstein, D., Häubl, G., & Weber, E. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, 23, 487–504

⁵ Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics*, 116, 1149–1187.

⁶ Malhotra, S., Cheriff, A. D., Gossey, J. T., Cole, C. L., Kaushal, R., & Ancker, J. S. (2016). Effects of an e-Prescribing interface redesign on rates of generic drug prescribing: exploiting default options. *Journal of the American Medical Informatics Association*, 23, 891–898.

⁷ Pichert, D., & Katsikopoulos, K. V. (2008). Green defaults: Information presentation and pro-environmental behaviour. *Journal of Environmental Psychology*, 28, 63–73.

⁸ Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, 302, 1338–1339.

⁹ Donorregister. Personal communication, June 1, 2017

¹⁰ August, J. G. (2013). Modern models of organ donation: Challenging increases of federal power to save lives. *Hastings Constitutional Law Quarterly*, 40, 339–422.

¹¹ Siminoff, L. A., & Mercer, M. B. (2001). Public policy, public opinion, and consent for organ donation. *Cambridge Quarterly of Healthcare Ethics*, 10, 377–386.

-
- ¹² Grice, H. P. (1975). Logic and conversation. In P. Cole & N. L. Morgan (Eds.), *Syntax and semantics: Speech acts* (Vol. 3, pp. 41–58). New York, NY: Academic Press.
- ¹³ Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. *Advances in Experimental Social Psychology*, *26*, 123-162.
- ¹⁴ Choi, J. J., Laibson, D., & Madrian, B. C. (2011). \$100 bills on the sidewalk: Suboptimal investment in 401(k) plans. *Review of Economics and Statistics*, *93*, 748-763.
- ¹⁵ Rhee, N. (2013). *The retirement savings crisis: Is it worse than we think?* National Institute on Retirement Security, June.
- ¹⁶ Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow: Using behavioral economics to increase employee saving. *Journal of Political Economy*, *112*, 164-187.
- ¹⁷ Beshears, J., Dai, H., Milkman, K. L., & Benartzi, S. (2016). Framing the future: The risks of pre-commitment nudges and potential of fresh start messaging. Working Paper.
- ¹⁸ McKenzie, C. R., & Nelson, J. D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin & Review*, *10*, 596-602.
- ¹⁹ Sher, S., & McKenzie, C. R. (2006). Information leakage from logically equivalent frames. *Cognition*, *101*, 467-494.
- ²⁰ Dinner, I., Johnson, E. J., Goldstein, D. G., & Liu, K. (2011). Partitioning default effects: why people choose not to choose. *Journal of Experimental Psychology: Applied*, *17*, 332-341.
- ²¹ McKenzie, C. R., Liersch, M. J., & Finkelstein, S. R. (2006). Recommendations implicit in policy defaults. *Psychological Science*, *17*, 414-420.
- ²² Tannenbaum, D., & Ditto, P. H. (2011). *Information asymmetries in default options*. Working Paper.
- ²³ Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *Quarterly Journal of Economics*, *116*, 1149-1187.
- ²⁴ Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2009). The importance of default options for retirement saving outcomes: Evidence from the United States. In J. Brown, J. B. Liebman, & D. A. Wise (Eds.). *Social Security Policy in a Changing Environment* (pp. 167-195). Chicago, IL: University of Chicago Press.
- ²⁵ Brown, J. R., Farrell, A. M., & Weisbenner, S. J. (2012). *The downside of defaults*. NBER Working Paper. Retrieved from <http://www.nber.org/aging/rrc/papers/orrc12-05.pdf>.
- ²⁶ Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *Quarterly Journal of Economics*, *116*, 1149-1187.
- ²⁷ Tannenbaum, D., & Ditto, P. H. (2011). *Information asymmetries in default options*. Working Paper.
- ²⁸ McKenzie, C. R., Liersch, M. J., & Finkelstein, S. R. (2006). Recommendations implicit in policy defaults. *Psychological Science*, *17*, 414-420.
- ²⁹ Tannenbaum, D., & Ditto, P. H. (2011). *Information asymmetries in default options*. Working Paper.
- ³⁰ Liersch, M. J., & McKenzie, C. R. M. (2009). *In defaults we trust*. Unpublished Manuscript.
- ³¹ Agnew, J. R., & Szykman, L. R. (2005). Asset allocation and information overload: The influence of information display, asset choice, and investor experience. *Journal of Behavioral Finance*, *6*, 57-70.
- ³² Brown, C. L., & Krishna, A. (2004). The skeptical shopper: A metacognitive account for the effects of default options on choice. *Journal of Consumer Research*, *31*, 529-539.
- ³³ Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes*, *39*, 84-97.
- ³⁴ Keys, B. J., & Wang, J. (2014). *Perverse nudges: Minimum payments and debt paydown in consumer credit cards*. Working Paper. Retrieved from https://www.economicdynamics.org/meetpapers/2014/paper_323.pdf.
- ³⁵ Navarro-Martinez, D., Salisbury, L. C., Lemon, K. N., Stewart, N., Matthews, W. J., & Harris, A. J. (2011). Minimum required payment and supplemental information disclosure effects on consumer debt repayment decisions. *Journal of Marketing Research*, *48*, S60-S77.
- ³⁶ Stewart, N. (2009). The cost of anchoring on credit-card minimum repayments. *Psychological Science*, *20*, 39-41.

-
- ³⁷ Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological Science, 12*, 391-396.
- ³⁸ Loschelder, D. D., Friese, M., Schaerer, M., & Galinsky, A. D. (2016). The too-much-precision effect: When and why precise anchors backfire with experts. *Psychological Science, 27*, 1573-1587.
- ³⁹ Goswami, I., & Urminsky, O. (2016). When should the ask be a nudge? The effect of default amounts on charitable donations. *Journal of Marketing Research, 53*, 829-846.
- ⁴⁰ Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2004). For better or for worse: Default effects and 401(k) savings behavior. In D. A. Wise (Ed.), *Perspectives on the Economics of Aging* (pp. 81-126). Chicago, IL: University of Chicago Press.
- ⁴¹ Valenzuela, A., & Raghurir, P. (2009). Position-based beliefs: The center-stage effect. *Journal of Consumer Psychology, 19*, 185-196.
- ⁴² Dayan, E., & Bar-Hillel, M. (2011). Nudge to nobesity II: Menu positions influence food orders. *Judgment and Decision Making, 6*, 333-342.
- ⁴³ Koppell, J. G., & Steen, J. A. (2004). The effects of ballot position on election outcomes. *Journal of Politics, 66*, 267-281.
- ⁴⁴ Bar-Hillel, M. (2015). Position effects in choice from simultaneous displays: A conundrum solved. *Perspectives on Psychological Science, 10*, 419-433.
- ⁴⁵ Valenzuela, A., & Raghurir, P. (2009). Position-based beliefs: The center-stage effect. *Journal of Consumer Psychology, 19*, 185-196.
- ⁴⁶ Fox, C. R., Ratner, R. K., & Lieb, D. S. (2005). How subjective grouping of options influences choice and allocation: Diversification bias and the phenomenon of partition dependence. *Journal of Experimental Psychology: General, 134*, 538-551.
- ⁴⁷ Tannenbaum, D., Doctor, J. N., Persell, S. D., Friedberg, M. W., Meeker, D., Friesema, E. M., ... & Fox, C. R. (2015). Nudging physician prescription decisions by partitioning the order set: results of a vignette-based study. *Journal of general internal medicine, 30*, 298.
- ⁴⁸ Fox, C. R., & Clemen, R. T. (2005). Subjective probability assessment in decision analysis: Partition dependence and bias toward the ignorance prior. *Management Science, 51*, 1417-1432.
- ⁴⁹ Sonnemann, U., Camerer, C. F., Fox, C. R., & Langer, T. (2013). How psychological framing affects economic market prices in the lab and field. *Proceedings of the National Academy of Sciences, 110*, 11779-11784.
- ⁵⁰ Langer, T., & Fox, C. R. (2005). *Biases in allocation under risk and uncertainty: Partition dependence, unit dependence, and procedure dependence*. Working Paper.
- ⁵¹ Tannenbaum, D., Fox, C. R., Goldstein, N. J. (2013). *Partitioning menu items to nudge single-item choice*. Working Paper. Retrieved from <https://davetannenbaum.github.io/documents/pdepend.pdf>
- ⁵² Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in Experimental Social Psychology, 24*, 201-234.
- ⁵³ Mulder, L. B. (2008). The difference between punishments and rewards in fostering moral concerns in social decision making. *Journal of Experimental Social Psychology, 44*, 1436-1443.
- ⁵⁴ Tannenbaum, D., Valasek, C. J., Knowles, E. D., & Ditto, P. H. (2013). Incentivizing wellness in the workplace: Sticks (not carrots) send stigmatizing signals. *Psychological Science, 24*, 1512-1522.
- ⁵⁵ Evers, E. R., Inbar, Y., Blanken, I., & Oosterwijk, L. D. (2016). When do people prefer carrots to sticks? A robust “matching effect” in policy evaluation. *Management Science*. Advance online publication.
- ⁵⁶ Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies, 70*, 489-520.
- ⁵⁷ Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *American Economic Review, 87*, 746-755.
- ⁵⁸ Executive Order No. 13676 (2014). *Combating antibiotic-resistant bacteria*. Retrieved from <https://www.whitehouse.gov/the-press-office/2014/09/18/executive-order-combating-antibiotic-resistant-bacteria>.

-
- ⁵⁹ Review on Antimicrobial Resistance (2016). *Tackling drug-resistant infections globally: Final report and recommendations*. Retrieved from <https://amr-review.org/Publications>.
- ⁶⁰ Meeker, D., Linder, J. A., Fox, C. R., Friedberg, M. W., Persell, S. D., Goldstein, N. J., Knight, T. K., Hay, J. W., & Doctor, J. N. (2016). Effect of behavioral interventions on inappropriate antibiotic prescribing among primary care practices: a randomized clinical trial. *JAMA*, *315*, 562-570.
- ⁶¹ Barnett, M. L., & Linder, J. A. (2014). Antibiotic prescribing for adults with acute bronchitis in the United States, 1996-2010. *JAMA*, *311*, 2020-2022.
- ⁶² This is not apparent from the published version of the paper which displays smoothed data, but it is in the raw data, available from the authors.
- ⁶³ Gosnell, G. K., List, J. A., & Metcalfe, R. (2016). *A new approach to an age-old problem: Solving externalities by incenting workers directly*. NBER Working Paper No. 22316.
- ⁶⁴ Schwartz, D., Fischhoff, B., Krishnamurti, T., & Sowell, F. (2013). The Hawthorne effect and energy awareness. *Proceedings of the National Academy of Sciences*, *110*, 15242-15246.
- ⁶⁵ Levitt, S. D., & List, J. A. (2011). Was there really a Hawthorne effect at the Hawthorne plant? An analysis of the original illumination experiments. *American Economic Journal: Applied Economics*, *3*, 224-238.
- ⁶⁶ Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*, 75-98.
- ⁶⁷ Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review*, *102*, 33-48.
- ⁶⁸ Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*, 255-275.
- ⁶⁹ Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, *21*, 1-31.
- ⁷⁰ Brehm, J. W. (1966). *A Theory of Psychological Reactance*. Oxford, UK: Academic Press.
- ⁷¹ Clee, M. A., & Wicklund, R. A. (1980). Consumer behavior and psychological reactance. *Journal of Consumer Research*, *6*, 389-405.
- ⁷² Wicklund, R. A. (1974). *Freedom and Reactance*. Oxford, UK: Lawrence Erlbaum.
- ⁷³ Sunstein, C. R. (2017). Nudges that fail. *Behavioural Public Policy*, *1*, 4-25.
- ⁷⁴ Clee, M. A., & Wicklund, R. A. (1980). Consumer behavior and psychological reactance. *Journal of Consumer Research*, *6*, 389-405.
- ⁷⁵ Jung, J. Y., & Data, B. A. M. (2016). American attitudes toward nudges. *Judgment and Decision making*, *11*, 62-74.
- ⁷⁶ Tannenbaum, D., & Ditto, P. H. (2011). *Information asymmetries in default options*. Working Paper.
- ⁷⁷ Agnew, J. R., Szykman, L. R., Utkus, S. P., & Young, J. A. (2012). Trust, plan knowledge and 401(k) savings behavior. *Journal of Pension Economics & Finance*, *11*, 1-20.
- ⁷⁸ Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, *302*, 1338-1339.
- ⁷⁹ Davidai, S., Gilovich, T., & Ross, L. D. (2012). The meaning of default options for potential organ donors. *Proceedings of the National Academy of Sciences*, (38), 15201-15205.
- ⁸⁰ Young, S. D., Monin, B., & Owens, D. (2009). Opt-out testing for stigmatized diseases: A social psychological approach to understanding the potential effect of recommendations for routine HIV testing. *Health Psychology*, *28*, 675-681.
- ⁸¹ Bénabou, R., & Tirole, J. (2006) Incentives and prosocial behavior. *American Economic Review*, *96*, 1652-1678.
- ⁸² Kamenica, E. (2012). Behavioral economics and psychology of incentives. *Annual Review of Economics*, *4*, 427-452.
- ⁸³ Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, *25*, 191-209.
- ⁸⁴ Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies*, *29*, 1-18.

-
- ⁸⁵ Chao, M. (2017). Demotivating incentives and motivation crowding out in charitable giving. *Proceedings of the National Academy of Sciences*. Advance online publication.
- ⁸⁶ Lacetera, N., & Macis, M. (2010). Do all material incentives for pro-social activities backfire? The response to cash and non-cash incentives for blood donations. *Journal of Economic Psychology*, *31*, 738-748.
- ⁸⁷ Lacetera, N., Macis, M., & Slonim, R. (2012). Will there be blood? Incentives and displacement effects in pro-social behavior. *American Economic Journal: Economic Policy*, *4*, 186-223.
- ⁸⁸ Lacetera, N., & Macis, M. (2008). *Motivating altruism: A field study*. IZA Discussion Paper 3770.
- ⁸⁹ Lacetera, N., & Macis, M. (2010). Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *Journal of Economic Behavior & Organization*, *76*, 225-237.
- ⁹⁰ Mellström, C., & Johannesson, M. (2008). Crowding out in blood donation: Was Titmuss right? *Journal of the European Economic Association*, *6*, 845-863.
- ⁹¹ Thaler, R. H., Sunstein, C. R., & Balz, J. P. (2012). Choice architecture. In E. Shafir (Ed.). *The Behavioral Foundations of Public Policy* (pp. 428-439). Princeton, NJ: Princeton University Press.
- ⁹² Jones, E. E., & Davis, K. E. (1965). From acts to dispositions the attribution process in person perception. *Advances in Experimental Social Psychology*, *2*, 219-266.
- ⁹³ Chaiken, S., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.). *Unintended Thought* (pp. 212-252). New York, NY: Guilford Press.
- ⁹⁴ Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In R. E. Petty (Ed.). *Communication and Persuasion* (pp. 1-24). New York, NY: Springer.
- ⁹⁵ Folkes, V. S. (1988). Recent attribution research in consumer behavior: A review and new directions. *Journal of Consumer Research*, *14*, 548-565.
- ⁹⁶ Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*, 548-573.
- ⁹⁷ Jones, E. E., & Davis, K. E. (1965). From acts to dispositions the attribution process in person perception. *Advances in Experimental Social Psychology*, *2*, 219-266.
- ⁹⁸ Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136-153.
- ⁹⁹ Agnew, J. R., & Szykman, L. R. (2005). Asset allocation and information overload: The influence of information display, asset choice, and investor experience. *Journal of Behavioral Finance*, *6*, 57-70.
- ¹⁰⁰ Hadar, L., Sood, S., & Fox, C. R. (2013). Subjective knowledge in consumer financial decisions. *Journal of Marketing Research*, *50*, 303-316.
- ¹⁰¹ Fox, C. R., & Weber, M. (2002). Ambiguity aversion, comparative ignorance, and decision context. *Organizational Behavior and Human Decision Processes*, *88*, 476-498.
- ¹⁰² Hadar, L., Tannenbaum, T., & Fox, C. R. (2017). *Subjective knowledge attenuates default effects*. Working Paper.
- ¹⁰³ Tannenbaum, D., Fox, C. R., & Rogers, T. (2017). On the misplaced politics of behavioral policy interventions. *Nature Human Behavior*. Advance online publication.
- ¹⁰⁴ Forehand, M. R., & Grier, S. (2003). When is honesty the best policy? The effect of stated company intent on consumer skepticism. *Journal of Consumer Psychology*, *13*, 349-356.
- ¹⁰⁵ Costa, D. L., & Kahn, M. E. (2013). Energy conservation “nudges” and environmentalist ideology: Evidence from a randomized residential electricity field experiment. *Journal of the European Economic Association*, *11*, 680-702.
- ¹⁰⁶ Reisch, L. A., & Sunstein, C. R. (2016). Do Europeans like nudges? *Judgment and Decision Making*, *11*, 310-325.
- ¹⁰⁷ Goldstein, D. G., Johnson, E. J., Herrmann, A., & Heitmann, M. (2008). Nudge your customers toward better choices. *Harvard Business Review*, *86*, 99-105.

¹⁰⁸ De Haan, T., & Linde, J. (2017). ‘Good nudge lullaby’: Choice architecture and default bias reinforcement. *The Economic Journal*. Advance online publication.

¹⁰⁹ Felsen, G., Castelo, N., & Reiner, P. B. (2013). Decisional enhancement and autonomy: public attitudes towards overt and covert nudges. *Judgment and Decision Making*, *8*, 202-213.

¹¹⁰ House of Lords (2011). *Behaviour Change*. Science and Technology Select Committee Report, HL paper 179.

¹¹¹ Loewenstein, G., Bryce, C., Hagmann, & Rajpal, S. (2015). *Behavioral Science & Policy*, *1*, 35-42.

¹¹² Steffel, M., Williams, E. F., & Pogacar, R. (2016). Ethically deployed defaults: Transparency and consumer protection through disclosure and preference articulation. *Journal of Marketing Research*, *53*, 865-880.

¹¹³ Wood, W., & Quinn, J. M. (2003). Forewarned and forearmed? Two meta-analysis syntheses of forewarnings of influence appeals. *Psychological Bulletin*, *129*, 119-138.

¹¹⁴ Campbell, M. C., Mohr, G., & Verlegh, P. W. (2012). Can disclosures lead consumers to resist covert persuasion? The important roles of disclosure timing and type of response. *Journal of Consumer Psychology*, *23*, 483-495.

¹¹⁵ Forehand, M. R., & Grier, S. (2003). When is honesty the best policy? The effect of stated company intent on consumer skepticism. *Journal of Consumer Psychology*, *13*, 349-356.

¹¹⁶ Steffel, M., Williams, E. F., & Pogacar, R. (2016). Ethically deployed defaults: Transparency and consumer protection through disclosure and preference articulation. *Journal of Marketing Research*, *53*, 865-880.