

Optimal nudging for cognitively bounded agents: A framework for modeling, predicting, and controlling the effects of choice architectures

Frederick Callaway^{1*}, Mathew Hardy^{1*}, and Thomas L. Griffiths^{1,2}

¹Department of Psychology, Princeton University, Princeton, NJ

²Department of Computer Science, Princeton University, Princeton, NJ

People’s judgments and decisions often deviate from classical notions of rationality, incurring costs both to themselves and to society. One way to reduce the costs of poor decisions is to redesign the decision problems people face to encourage better choices. While often subtle, these *nudges* can have dramatic effects on behavior and are increasingly popular in public policy, healthcare, and marketing. Although nudges are often designed with psychological theories in mind, they are typically not formalized in computational terms and their effects can be hard to predict. As a result, designing nudges can be difficult and time-consuming. To address this challenge, we propose a computational framework for understanding and predicting the effects of nudges. Our framework builds on recent work modeling human decision-making as adaptive use of limited cognitive resources, an approach called *resource-rational analysis*. Concretely, nudges change the optimal sequence of cognitive operations an agent should execute, which in turn influences the agent’s behavior. We first show that our framework can account for known effects of nudges based on default options, suggested alternatives, and information highlighting. In each case, we validate the model’s predictions in an experimental process-tracing paradigm. We then show how the framework can be used to automatically construct optimal nudges, and demonstrate that these nudges improve people’s decisions more than intuitive heuristic approaches. Overall, our results show that resource-rational analysis is a promising framework for formally characterizing and constructing nudges.

Keywords: Nudging, choice architecture, resource-rational analysis

How do people choose when to recycle, how much to save, or what to buy for lunch? When facing decisions both large and small, people’s choices are often at odds with classical notions of rationality. These deviations are not only theoretically important, but can also incur large costs for both individuals and societies (Kahneman, Slovic, & Tversky, 1982). For example, a majority of Americans report undersaving for retirement (Choi et al., 2006), which may be partly due to people’s inconsistent time preferences, tendency to neglect compounding, and bias towards procrastination (Goda, Levy, Manchester, Sojourner, & Tasoff, 2015; Thaler & Benartzi, 2004).

In an effort to reduce these costs, researchers have proposed using theoretical and empirical results from psychology to redesign the *choice architecture*, or way that decisions are structured and framed, to help people make better choices

(Thaler & Sunstein, 2008). For example, retirement savings can be increased by having employees opt out of automatic contributions, rather than opting in (Madrian & Shea, 2001). These *nudges* are an increasingly popular complement to traditional interventions such as educational programs, legislation, and tax incentives, and are often significantly less expensive to administer (Benartzi et al., 2017).

While promising, there is no widely accepted formal framework for modeling the effects of different choice architectures on behavior. This introduces three substantial challenges for nudge theory. First, models of nudges are often domain-specific and ad hoc (Chetty, 2015; Hausman & Welch, 2010; Kusters & Van der Heijden, 2015; Willis, 2013; Yeung, 2012). This makes it difficult to make robust predictions in new settings (Kusters & Van der Heijden, 2015; Moseley & Stoker, 2013; Yeung, 2012). Second, there is often disagreement about what behaviors nudges should aim to promote (Goodwin, 2012; Tannenbaum, Fox, & Rogers, 2017). When choice architects and decision makers have different goals, nudges could be seen as manipulative or even exploitative (Hausman & Welch, 2010; Wilkinson, 2013). Identifying appropriate goals for nudges is especially challenging in populations with heterogeneous preferences and needs (Carroll, Choi, Laibson, Madrian, & Metrick, 2009;

* FC and MH contributed equally to this work. Correspondence concerning this article should be addressed to Frederick Callaway <fredcallaway@princeton.edu>. All code and data used to generate the results in this paper can be found at <https://github.com/fredcallaway/optimal-nudging>

Mills, 2020). Third, developing new nudges often involves an iterative process of search and experimentation (Moseley & Stoker, 2013; Vlaev, King, Dolan, & Darzi, 2016). This process is slow and expensive, and it constrains nudges to choice architectures that researchers find intuitive.

In this paper, we propose using resource-rational analysis as a formal framework for modeling nudges and predicting their effects. Resource-rational analysis is an approach to deriving models of human behavior by assuming that people make optimal use of their limited computational resources (Griffiths, Lieder, & Goodman, 2015; Lieder & Griffiths, 2020; c.f. Gershman, Horvitz, & Tenenbaum, 2015; Howes, Lewis, & Vera, 2009; Lewis, Howes, & Singh, 2014; Sims, 2003). That is, while classical rational models assume that people are rational with respect to their choices, resource-rational models assume that people are rational with respect to how much they think and what they think about. Importantly, which factors are worth considering depends critically on the ease with which they can be considered. This suggests a principled way of understanding nudging: by modifying the accessibility of different pieces of information, nudges change what information a decision maker considers, which in turn influences their choices.

The framework of resource-rational analysis can help address the three challenges facing choice architects described above. First and foremost, it provides a way to build mathematically explicit models of how a person's decision-making process—and ultimately, their decision—depends on aspects of the choice architecture. This not only makes it possible to provide rigorous formal explanations for the effects of nudges, but also to *predict* how those effects will change under new decision contexts or variations of the nudge. Second, by formalizing decision-making as a cost-benefit optimization problem, resource-rational analysis provides a new way to conceptualize and quantify the *goals* of nudges. In particular, we can evaluate a nudge not just based on the effect it has on a person's decision, but also based on the effect it has on their deliberation. Finally, building on the first two contributions, we present an automated method for constructing *optimal* nudges, in which we use optimization algorithms to identify the choice architecture that best satisfies a given goal.

The paper is organized as follows. We first provide a brief review of nudging and resource rational analysis. We then present our formal framework for modeling nudging and introduce an experimental paradigm for testing our predictions. We apply our approach to three popular nudges—default options, suggestions, and information highlighting. For each nudge, we first use resource-rational analysis to model its effects, and then test predictions from our model in a behavioral experiment. After providing a formal account of these nudges, we show how our framework can be used to automatically construct *optimal* information highlighting nudges.

We test our approach to optimal information highlighting in two experiments, comparing nudges determined by our procedure with those constructed by a heuristic and those constructed randomly. We conclude by discussing the implications of our findings and approach to nudge theory, ways in which our framework could be improved, and promising areas for future research.

Nudging

Nudging is an approach to improving people's choices by changing the way decisions are framed and presented (Sunstein, 2019; Thaler & Sunstein, 2008). Nudges use findings from psychology and behavioral economics to change the structure of a decision without restricting people's freedom of choice or changing their economic incentives. While often simple, nudges have been effective in a diverse set of domains, including education (Damgaard & Nielsen, 2018), finance (Cai, 2020), healthcare (Voyer, 2015), energy consumption (Lehner, Mont, & Heiskanen, 2016), and tax compliance (Antinyan & Asatryan, 2019), among others (Sunstein, 2016; Szaszi, Palinkas, Palfi, Szollosi, & Aczel, 2018).

Most research on nudging has been guided by research on heuristics and biases (Kahneman et al., 1982). In this paradigm, researchers design choice architectures that either attempt to mitigate the effect of a bias or instead leverage the bias to guide behavior. Consider, for example, the widely studied tendency for people to procrastinate performing important tasks (Steel, 2007). The effects of this bias can be reduced by giving people tighter deadlines (O'Donoghue & Rabin, 1998); for example, individuals redeem more gift certificates when they expire more quickly (Shu & Gneezy, 2010). On the other hand, people's bias towards procrastination can be used as a tool for improving choice. For example, Thaler and Benartzi (2004) show that allowing employees to commit to saving a proportion of future salary increases can improve savings rates. As we review further below, researchers have catalogued many examples of effective nudges, and provided plausible cognitive mechanisms underlying their effects. However, these explanations have typically not been formalized mathematically, and they are often highly specific to each individual case.

A more recent line of work has thus focused on developing formal models that capture of a wide range of nudges (Felsen & Reiner, 2015; Lin, Osman, & Ashcroft, 2017). These models typically abstract away from the specific psychological mechanisms involved, and instead characterize nudges using the formal tools of economics. To account for the effect of nudges, which by definition do not substantially change economic incentives, these models employ a distinction between a decision maker's decision utility (the objective they maximize when making a decision), and their experienced utility (the well-being they derive from their choice; Kahneman, Wakker, & Sarin, 1997). Specifically, nudges are modeled

as changes to the decision utility or the perceived cost of a good. These models of nudges have allowed economists to incorporate nudging into analyses of welfare and taxation (Allcott & Kessler, 2019; Carlsson & Johansson-Stenman, 2019; Chetty, 2015; Farhi & Gabaix, 2020). In a recent extension of this approach, Löfgren and Nordblom (2020) consider how properties of a decision determine when people make heuristic (and therefore, nudgeable) choices. However, because these models abstract away from specific choice architectures (summarizing them as simple scalar utility offsets), they are unable to make specific predictions about how different nudges will affect people's decisions.

Thus, one line of work has produced intuitive, mechanistic accounts of individual nudges, while another has produced formal, abstract accounts of nudging as a whole. However, to the best of our knowledge, no previous work has proposed a formal framework that can make predictions about the effects of specific choice architectures. In this paper, we propose that resource-rational analysis provides such a framework. We demonstrate the approach using three commonly-used nudges as case studies: default options, suggestions, and information highlighting. We now describe each type of nudge in detail and briefly review findings on their impact on choice.

Default options

Perhaps the best known and most successful type of nudge involves reconfiguring or introducing default options. Default options are those that the decision maker (hereafter, the *agent*) will select if they do not act. That is, default options are chosen when no decision is made, and thus define the status quo. For example, in the United States people must sign up to be organ donors. In other countries, however, the default is switched: people are considered organ donors unless they request not to be.

While subtle, defaults can have substantial effects on people's choices in both field (Bergeron, Doyon, Saulais, & Labrecque, 2019; Momen & Stoerk, 2014) and laboratory (Huh, Vosgerau, & Morewedge, 2014) settings. For example, organ donation rates in countries with opt-out programs have significantly higher donation rates than countries with opt-in programs (Abadie & Gay, 2006; Johnson & Goldstein, 2004). Similarly, defaults have been shown to have large effects on a wide range of decisions, such as investment and saving (Madrian & Shea, 2001), insurance selection (Johnson, Hershey, Meszaros, & Kunreuther, 1993), and charitable donations (Goswami & Urminsky, 2016).

Despite their widespread use and overall success, defaults are not always effective (Sunstein, 2017). A recent meta-analysis found several studies in which default options had no significant effect on choice, and there was considerable variation in the effect size among those with significant effects (Jachimowicz, Duncan, Weber, & Johnson, 2019). Ex-

plaining why defaults work when they do—and better yet, predicting new contexts in which they will be effective—is thus an important goal in nudging research.

Several explanations for the influence of defaults have been offered. First, making a choice is costly—an agent may decide that evaluating possible alternatives is simply not worth their effort when a default is offered (Johnson & Goldstein, 2003; Johnson et al., 2012). This is supported by research showing that placing people under cognitive load (Huh et al., 2014) or time pressure (White, Jiang, & Albaracin, 2021) increases the chance that they stick with the default. Second, the agent may interpret the default as an implicit recommendation or endorsement from the choice architect or policy maker (Gigerenzer, 2008; Johnson & Goldstein, 2003; McKenzie, Liersch, & Finkelstein, 2006). This may cause defaults to be less effective in domains where people have expertise (Brown, Farrell, & Weisbenner, 2012; Löfgren, Martinsson, Hennlock, & Sterner, 2012) or perceive the choice architect as having interests differing from their own (Tannenbaum et al., 2017). Third, if the default option is used as a reference point, loss aversion may bias the agent towards the status quo, or to sticking with the default (Dinner, Johnson, Goldstein, & Liu, 2011; Fryer Jr, Levitt, List, & Sadoff, 2012). Indeed, each of these effects may influence different people on the same decision problem (Brown et al., 2012).

Suggested alternatives

In many domains, people are offered a suggestion before or after making a choice. For example, digital recommender systems allow consumers to identify options they may not have considered on their own (Schafer, Konstan, & Riedl, 1999), and salespeople often “upsell”, suggesting lucrative alternatives or additions after a customer's initial choice. Research has shown that when used for the consumer's benefit, suggestions can be effective instruments for improving people's choices (van Kleef, van den Broek, & van Trijp, 2015). For example, in a twist on the well-known “supersizing” upsell, Schwartz, Riis, Elbel, and Ariely (2012) showed that up to 33% of customers at a fast food restaurant accepted an offer to downsize a side order, significantly reducing their calorie consumption. Similarly, suggesting nutritious side dishes can increase healthy purchases (Vercellis, 2009), and offering restaurant customers the opportunity to “wrap” their leftovers after a meal may reduce food waste (Hamerman, Rudell, & Martins, 2018). The effectiveness of suggested alternatives is not limited to food choice—Forget, Chiasson, Van Oorschot, and Biddle (2008) show that providing more secure alternatives to user-generated passwords can significantly improve password security.

Suggestions can also be given before an initial choice, an especially common approach in digital recommendation systems (Jesse & Jannach, 2021; Xiao & Benbasat, 2007). In-

deed, suggestions given early in the deliberation process may be especially effective in impacting choice (Forwood, Ahern, Marteau, & Jebb, 2015). For example, Bothos, Apostolou, and Mentzas (2015) show that a recommender system can help commuters identify alternative eco-friendly routes they would not have considered on their own.

While often successful, it remains unclear why and how upsells and other suggestions influence people's choices. Suggestions may break the automatic behavior, or "script," of certain situations, allowing people to exert more self-control and make better choices than they otherwise would (Schwartz et al., 2012). In other domains, external suggestions may allow people to justify norm-violating choices or behaviors (Hamerman et al., 2018). Lastly, because suggestions are often accompanied by novel information highlighting their attractiveness (Heidig, Wentzel, Tomczak, Wiecek, & Falzl, 2017), people may tend to systematically overestimate their quality.

Information highlighting

Default options and suggestions influence people's choices by making certain choices easier than others or providing additional information about certain options. Information highlighting nudges, by contrast, influence people's behavior by modifying the presentation of choice-relevant information. By making certain information more or less salient, information highlighting nudges take advantage of the limited attention, effort, and time people have to make their decision, and often influence people's choices without their knowledge.

While common in many domains, information highlighting is especially popular in designing "foodscapes"—the physical and digital environments where people purchase or consume food—to help people make healthier choices (Bucher et al., 2016; Elswiler, Trattner, & Harvey, 2017; Starke, Kløverød Brynstad, Hauge, & Løkeland, 2021; Wilson, Buckley, Buckley, & Bogomolova, 2016). One popular approach to promoting healthier eating choices is to manipulate the labeling of nutritional information on food packaging. However, consumers often attend to just one or two product features when deciding what to buy (Kahnemann, Tversky, & Rogers, 2013), and so the effects of labeling depend critically on how information is displayed, not just what information is available (Liu, Wisdom, Roberto, Liu, & Ubel, 2014). Indeed, simply providing additional information about each option (such as the number of calories it contains) often has little, if any, effect on choice (Kiszko, Martinez, Abrams, & Elbel, 2014; Loewenstein, Asch, Friedman, Melichar, & Volpp, 2012; Sinclair, Cooper, & Mansfield, 2014) as many consumers simply ignore the information (Krukowski, Harvey-Berino, Kolodinsky, Narsana, & DeSisto, 2006). Instead, effective information highlighting typically involves summarizing a limited number of features

with simple visual cues (Lin et al., 2017). These effects, of course, have long been exploited in marketing contexts, where product design, packaging, and labeling is manipulated to increase sales (Deliza & MacFie, 2001).

For example, Ecuador recently introduced mandatory "traffic light" labeling of food items' fat, sugar, and salt content. In these labels, the concentration of each nutrient is listed at one of four levels, with each level represented by a unique color—none (white), low (green), medium (orange), and high (red). Research has found that these labels have improved healthy eating behavior in Ecuador (Sandoval, Carpio, & Sanchez-Plata, 2019), mirroring findings on the effects of similar labels in restaurants and cafeterias in other countries (Ellison, Lusk, & Davis, 2014; Sonnenberg et al., 2013; Thorndike, Riis, Sonnenberg, & Levy, 2014). However, the impact of traffic light labeling appears to be highly dependent on individual characteristics such as age (Freire, Waters, & Rivas-Mariño, 2017), health attitudes (Freire, Waters, Rivas-Mariño, Nguyen, & Rivas, 2017), and socioeconomic status (Orozco, Ochoa, Muquinche, Padro, & Melby, 2017; Sandoval et al., 2019).

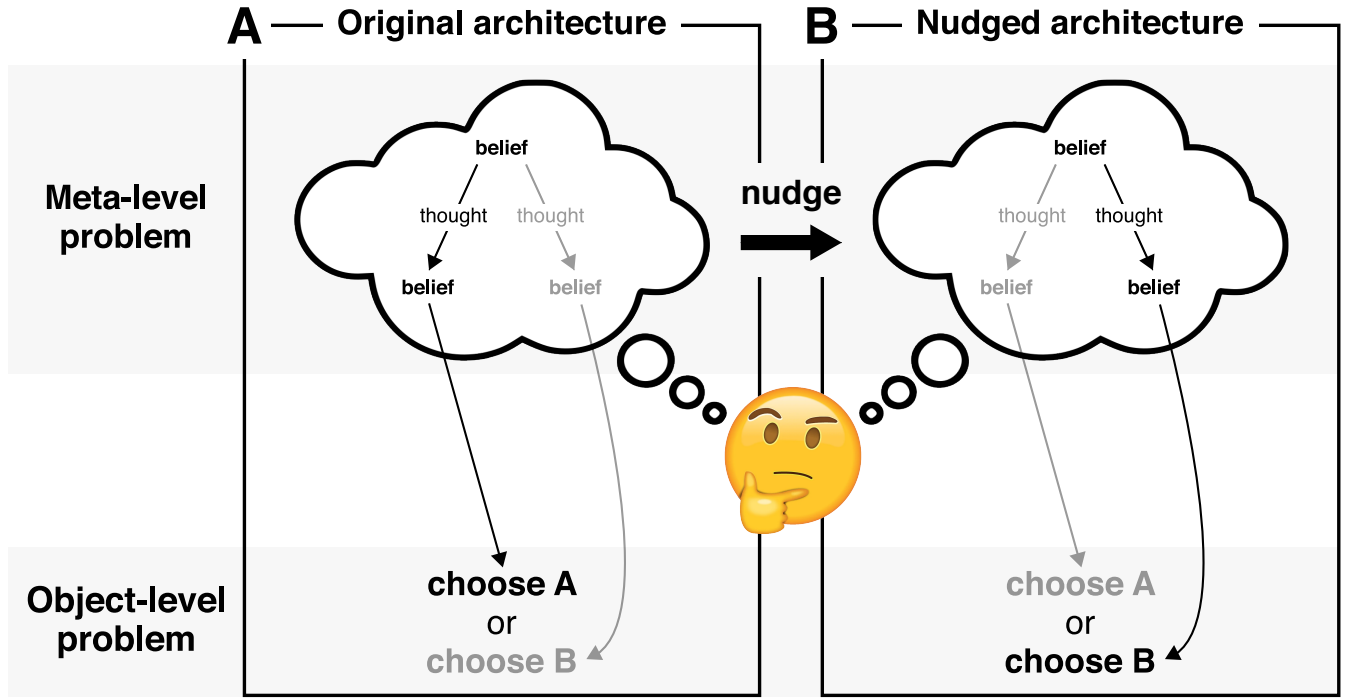
Summary

Nudges are changes to the structure or framing of a decision that do not restrict people's freedom of choice or change their economic incentives. While nudging is a promising tool for improving choice and reducing bias, there is no unifying theoretical framework for modeling the effects of different choice architectures on deliberation and behavior (Buckley, 2009; Yeung, 2012). In this paper, we argue that resource-rational analysis is a promising framework for doing so.

A resource-rational approach to nudging

Resource-rational analysis is a formal framework for deriving cognitive models based on the assumption that people act optimally with respect to their limited cognitive resources. Within this approach, a cognitive process is understood as the solution to an optimization problem, where the objective function explicitly trades off external utility with internal computational cost. Given that exactly solving complex problems generally incurs high computational cost, resource-rational models rarely predict that people will even attempt to identify exact solutions. Instead, they find solutions that balance choice quality with effort. As a result, resource-rational models often make behavioral predictions that differ dramatically from classical theories of rationality, and have been shown to account for a wide range of apparent biases and errors in human decision-making (see Bhui, Lai, and Gershman (2021); Lieder and Griffiths (2020) for reviews).

A key intuition behind resource-rational analysis is that the effectiveness of a decision-making strategy depends not only on the agent's external environment, but also on their

Figure 1*A resource rational approach to nudging*

Note. (A) We model decision-making as a process in which an agent executes a sequence of cognitive operations (thoughts) that update their beliefs about the value of the available choices. The space of possible thoughts and beliefs defines the *meta-level* problem of decision-making: which factors should the agent consider, and which should they ignore? These meta-level decisions determine the belief upon which the agent ultimately makes a choice, and therefore the choice itself. (B) We model nudges as modifications to the meta-level problem. For example, a nudge might make some thoughts easier to have than others, leading the agent to arrive at a different belief about the value of each option. Critically, this can change their ultimate choice without changing anything about the options themselves (that is, without modifying the *object-level* problem).

internal environment, that is, the workings of the agent’s own mind (c.f. Simon, 1955). How can we formally characterize that internal environment? Recent work has done so using the same tools typically used to model external environments. In this approach, cognitive processes (such as decision-making) are modeled as *sequential decision problems* (Callaway, Lieder, et al., 2018; Callaway, Rangel, & Griffiths, 2021; Chen, Chang, & Howes, 2021; Lieder, Krueger, & Griffiths, 2017). But while a traditional sequential decision problem is defined by states of the world and physical actions, the internal—or *meta-level*—decision problem is defined by mental states and cognitive operations. Intuitively, making a decision often involves considering multiple factors, and at each moment we have some control over which factor we will consider next. These meta-level decisions about what to think about determine our future beliefs, which in turn determine the external—or *object-level*—choices we ultimately make.

As illustrated in Figure 1, formalizing decision-making

as a sequential decision problem provides a new way to understand the effects of nudges. Rather than changing incentives or limiting choice (changes to the object-level problem), nudges make it easier to consider some factors than others or change the information the decision-maker starts with. That is, they change the meta-level problem. By changing which thoughts a decision-maker is likely to have, nudges indirectly change the beliefs upon which a choice is made, potentially changing the choice itself.

Modeling nudges in this way has three consequences.¹ First, resource-rational models allow us to make quantitative predictions about the effects of different nudges. Second, formalizing the computational cost of decision-making allows us to formalize different possible goals for nudges, such as making decisions easier. Finally, building on the first two contributions, we can use resource-rational models to con-

¹For a related discussion about the value of behavioral economics for public policy, see Chetty (2015).

struct *optimal* nudges, that is, nudges that best accomplish their goals. We now briefly expand on each contribution.

Predicting the effects of nudges

A key insight from previous research on nudging is that we can use psychological theory to guide the construction of choice architectures. However, rather than originating from a unified formal framework, this guidance is typically domain-specific and ad hoc (Chetty, 2015; Yeung, 2012; for similar arguments about behavioral law and economics, see Rostain, 2000). In contrast, if we can specify a mathematically explicit model of how a person's decision-making process depends on different aspects of the choice architecture, we can make quantitative predictions about the effects of different nudges across a range of domains. Importantly, the optimality assumption underlying resource-rational analysis allows us to make these predictions *a priori*. Indeed, all of the models we report in this paper have no free parameters. This contrasts with the standard cognitive modeling approach, where experimental data is often necessary to fit the parameters that govern the relationship between stimulus and response (in this case, choice architecture and choice).

Specifying the goals of nudges

Most previous applications of nudging have set the goal of encouraging people to make a specific choice, typically the one the choice architect assumes to be the best. However, this goal is often not made explicit to consumers, and does not take into account the possibility that the consumer may have different values than the choice architect. As a result, nudges have been criticized for their paternalism and lack of transparency (Goodwin, 2012; Wilkinson, 2013). Indeed, surveys indicate that people sometimes find nudges to be manipulative or unethical (Jung & Mellers, 2016).

An alternative approach is to design nudges with the explicit goal of increasing the consumer's utility, i.e., helping them make the best choice for themselves. However, by focusing solely on the choices people make, we ignore another positive effect nudges can have: They can make decisions *easier*. For example, consider the choice of which of two hotels to book as a conference venue, where one hotel has the unfortunate property of being unavailable on the chosen dates. Highlighting this information early on will not change the booker's final decision, but it could prevent unnecessary effort evaluating the relative merits of each hotel's coffee service. Resource-rational analysis formalizes the cost of that effort in a way that allows a choice architect to directly balance between the ease of decision-making and the possibility of making a suboptimal choice.

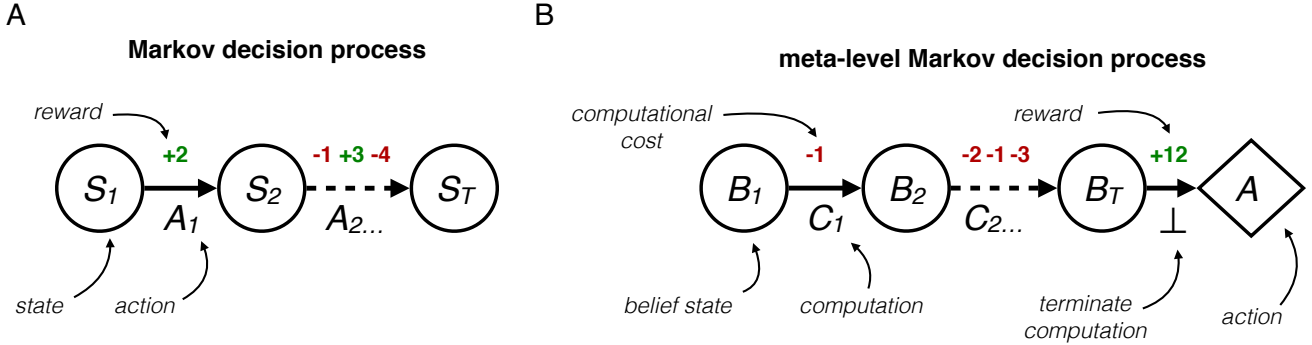
Constructing optimal nudges

Designing nudges is a challenging task. Because there is little formal theory about *why* some nudges are effective and others and not (Yeung, 2012), developing nudges often involves an iterative, experimental approach that can be time-consuming and expensive. This problem looms especially large when the space of possibilities is large: For example, when designing information highlighting nudges, choice architects not only have to choose the goal of the nudge, but also which type of modifications to make. These decisions can involve complex interactions and tradeoffs—should good options be made more appealing, or bad options less so? Should decisions be designed to encourage any good choice, or only the best one? And how can choice architects ensure that the nudge is beneficial to people with idiosyncratic preferences?

To address these challenges, we propose an automated method for constructing *optimal* nudges based on our resource-rational framework. This contribution builds on the previous two. First, because resource-rational models can make strong quantitative predictions *a priori* (that is, without fitting parameters to data), we can predict effect of a candidate nudge without running an experiment. Second, because resource-rational analysis formalizes the cost-benefit tradeoff underlying resource-constrained decision-making, we can define mathematically precise goals that balance the ease and quality of the consumer's decisions. Together, these two pieces allow us to define an objective function that takes as input a candidate nudge and returns a scalar capturing the degree to which the nudge satisfies the choice architect's goal. This in turn makes it possible to apply optimization algorithms to design nudges that optimize this objective function. Critically, this whole procedure can be performed automatically, without supervision or human data.

Summary

The key insight in resource-rational analysis is to view rationality as a property of a decision-making *process*, rather than as a property of decisions themselves. This conception of rationality has three theoretical and practical advantages for the study and development of nudges. First, it provides us with a set of tools for modeling how factors other than utility influence people's choices; this allows us to predict the effect of different choice architectures. Second, it establishes a broader notion of utility, one that accounts for the difficulty of making a decision as well as the value of the chosen option; this allows us to specify objectives for nudges that target a more holistic notion of well-being. Third, drawing on the previous two advances, we can formalize the design of choice architecture as an optimization problem; this allows us to automatically construct nudges that accomplish arbitrary goals.

Figure 2*Formal framework: meta-level Markov decision processes*

Note. (A) A Markov decision process formalizes the problem of acting adaptively in a dynamic environment. The agent executes actions that change the state of the world and generate rewards, which the agent seeks to maximize. (B) A meta-level Markov decision process formalizes the problem of *deciding how to act* when computational resources are limited. The agent executes computations that update their belief state and incur computational cost. When the agent executes the termination operation $\perp$, they take an external action based on their current belief state.

In the following section, we provide an overview of the general framework and describe a specific model that instantiates the framework in the context of multi-attribute choice. In the remainder of the paper, we will show how this model can be applied to explain, predict, and optimize the effects of nudges.

Formal framework

As outlined above, the key insight underlying our framework is that cognitive processes such as decision-making can themselves be viewed as sequential decision problems. Drawing on a subfield of artificial intelligence known as *rational metareasoning* (Matheson, 1968; Russell & Wefald, 1991), we formalize this insight using the framework of *meta-level Markov decision processes* (meta-level MDPs; Hay, Russell, Tolpin, & Shimony, 2012). In this framework, a cognitive process is formalized as a sequential process of executing computational actions that update an agent’s beliefs about the world. At each moment, the agent must choose whether to continue deliberating, refining their beliefs but accruing computational cost, or to instead stop computing and make a decision. In the former case, they must additionally decide which computation to execute next (i.e., what to think about); in the latter case, they select the optimal action given their current belief and receive a reward associated with the external utility of that action.

In this section, we present the formal framework and show how it can be applied to multi-attribute choice, the domain in which we conduct our experimental case studies. We provide a non-technical summary at the end of this section.

Markov decision processes

The core mathematical object underlying our approach is the Markov decision process (MDP), illustrated in Figure 2A. MDPs are the standard formalism for modeling the sequential interaction between an agent and a stochastic environment. An MDP is defined by a set of states $\mathcal{S}$, a set of actions $\mathcal{A}$, a transition function T , and a reward function r . A state $s \in \mathcal{S}$ specifies the relevant state of the world. An action $a \in \mathcal{A}$ is an action the agent can perform. The transition function T encodes the dynamics of the world as a distribution of possible future states for each possible previous state and action. Finally, the reward function r specifies the reward or utility for executing a given action in a given state.

The standard goal in an MDP is to maximize the expected cumulative reward attained, that is, the *return*. Importantly, this may require incurring immediate losses (negative rewards) in order to get to a state from which a highly rewarding action can be executed. It is typically assumed that the agent selects their actions based on the current state; the mapping from state to action is called a policy, denoted π . Solving an MDP amounts to finding a policy that maximizes the expected return, that is, a mapping from states to actions that, when followed, maximizes the total reward one will receive on average.

In addition to their foundational role in artificial intelligence (Sutton & Barto, 2018), MDPs are widely used in models of human decision-making (Dayan & Daw, 2008). MDPs are the formal foundation for models of reinforcement learning (Niv, 2009) and model-based planning (Botvinick & Toussaint, 2012; Huys et al., 2015), as well as competition between the two systems (Daw, Niv, & Dayan, 2005; Karamati, Dezfouli, & Piray, 2011; Kool, Gershman, & Cush-

man, 2017). They have also been used to study information-seeking (Gottlieb, Oudeyer, Lopes, & Baranes, 2013; Hunt, Rutledge, Malalasekera, Kennerley, & Dolan, 2016), generalization (Tomov, Schulz, & Gershman, 2021), and hierarchical abstraction (Botvinick, Niv, & Barto, 2009; Solway et al., 2014). However, with a few notable exceptions (Dayan & Huys, 2008; Drugowitsch, Moreno-Bote, Churchland, Shadlen, & Pouget, 2012; Tajima, Drugowitsch, & Pouget, 2016), MDPs have primarily been used to model the sequential decision problems posed by the external world. In the following section, we show how this framework can be applied to model the sequential decision problem posed by one’s own cognitive architecture.

Meta-level Markov decision processes

Meta-level Markov decision processes (meta-level MDPs) extend the standard MDP formalism to model the sequential decision problem posed by resource-bounded computation (Hay et al., 2012). Like a standard MDP, there is a set of states $\mathcal{S}$, a set of actions $\mathcal{A}$, and a reward function r_{object} (we omit the transition function because we focus on one-shot decisions). These define the *object-level* problem: the external problem the agent must solve in the world. Additionally, the meta-level MDP defines a set of beliefs $\mathcal{B}$, a set of computations $\mathcal{C}$, and meta-level transition and reward functions, T_{meta} and r_{meta} . These define the *meta-level* problem: how to allocate limited computational resources in the service of solving the object-level problem.

As illustrated in Figure 2B, the meta-level problem is itself a sequential decision problem, analogous to one defined by a standard MDP. However, in the meta-level problem, the states are replaced by beliefs (mental states) and the actions are replaced by computations (cognitive operations). The meta-level transition function describes how computations update beliefs, and the meta-level reward function captures both computational cost and the object-level reward of the action that is ultimately executed. We provide a general formal definition of meta-level MDPs in Appendix A. In the next section, we define a specific meta-level MDP for multi-attribute choice.

A meta-level MDP for multi-attribute choice

In a multi-attribute choice problem, an agent must choose one out of a set of options that differ on a number of features, each of which the agent values to varying degrees. For example, consider the problem of purchasing a car; there are many options available, each of which has features such as price, fuel efficiency, comfort, and horsepower. The overall utility associated with purchasing each car depends on all of these features; thus, to make the best possible choice, one would need to consider all of them for every vehicle one could buy. However, the computational cost of exhaustively evaluating every car on the market makes such a strategy impractical—

indeed—irrational. Instead, a shrewd consumer would consider only a subset of all the possible options and features, first evaluating a large set of candidates on the most important features and only carefully evaluating the top contenders. Here, we formalize this kind decision problem as a meta-level MDP.

We begin with the object-level problem, that is, the decision posed by the world. We assume that the agent must select one from a set of options that vary on a number of features. The utility of each option is a linear combination of those features, with the weights capturing the agent’s preferences. We now formalize the object-level problem in terms of states, actions, and a reward function.

States. A state $s \in \mathcal{S}$ specifies the features of the options available to the agent, as well as the agent’s preferences for those features. Concretely, a state is defined by a pair $(X, \mathbf{w})$, where X is a matrix such that $x_{a,f}$ is the value of feature f for option a , and $\mathbf{w}$ is a vector such that w_f is the weight the agent puts on feature f . Critically, the agent does not have direct access to the state, but instead maintains a belief about the state as detailed below.

Actions. Each action $a \in \mathcal{A}$ selects one of the available options. There is one action for each option.

Object-level reward. The object-level reward for each action is simply the utility of the chosen option. We assume that this utility is a linear combination of the option’s features:

$$r_{\text{object}}(s, a) = \sum_f w_f x_{a,f}. \quad (1)$$

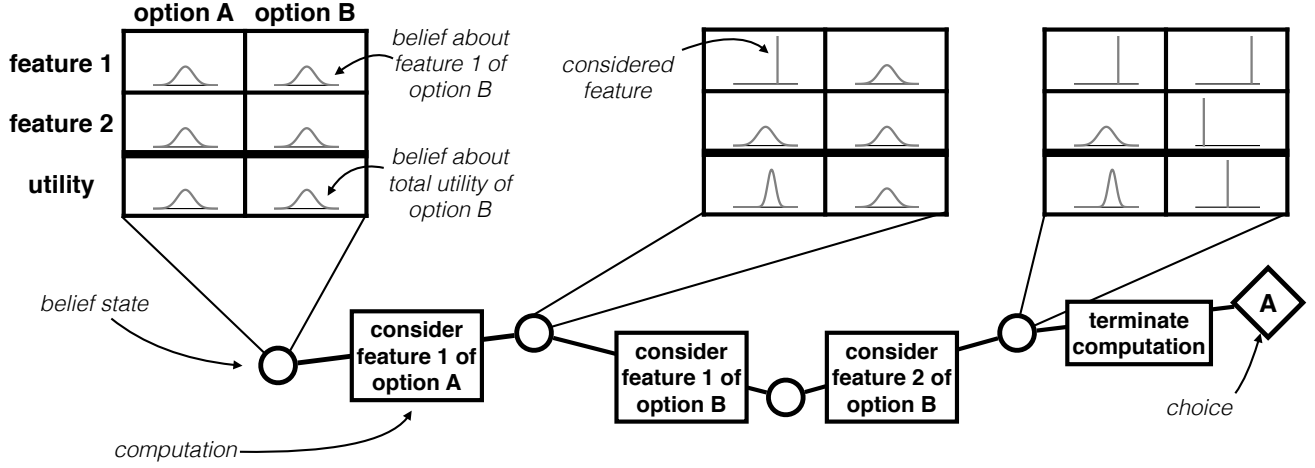
Note that the w and x on the right-hand side of the equation refer to the elements of s on the left-hand side.

We now formalize the meta-level problem, that is, the problem of how to decide which action to choose. Characterizing the precise computational architecture underlying human multi-attribute choice is an active research area (Berkowitsch, Scheibehenne, & Rieskamp, 2014; Bhatia & Stewart, 2018; Busemeyer, Gluth, Rieskamp, & Turner, 2019; Cohen, Kang, & Leise, 2017; Howes, Warren, Farmer, El-Deredy, & Lewis, 2016; Noguchi & Stewart, 2018; Roe, Busemeyer, & Townsend, 2001; Ronayne & Brown, 2017; Russo & Doshier, 1983; Trueblood, Brown, & Heathcote, 2014; Usher & McClelland, 2004). For simplicity (and consistency with the experimental paradigm that we employ), we will use a highly simplified architecture in which the belief state simply captures whether each feature has been considered, assuming that all considered features are perfectly integrated into the subjective expected utility. Importantly, this is not intended to be an accurate characterization of the computations involved in naturalistic multi-attribute choice. Instead, we make these assumptions to illustrate the framework in a simplified setting.

Beliefs. A belief $b \in \mathcal{B}$ encodes the agent’s knowledge about the feature values and weights. Formally, it is a distribution over states. We assume that the agent knows their own

Figure 3

A meta-level MDP for multi-attribute choice



Note. The agent must select among a set of options that vary on a number of features that determine the total utility of each option (for this example, we assume equal weighting). Making this decision can itself be modeled as a sequential decision problem in which the agent updates their beliefs by executing mental operations. Each mental operation evaluates one feature of one option and updates the belief about the option's utility accordingly. At some point, the agent stops deliberating and selects the option that is best according to their belief.

preferences; the belief thus encodes the true weight for each feature, $\mathbf{w}$. However, they are uncertain about the features of each options, and thus uncertain about the overall utility of each option. As illustrated in Figure 3, we capture this uncertainty with a set of independent normal distributions for each feature value, such that the belief about feature $x_{a,f}$ is defined

$$b(x_{a,f}) \sim \text{Normal}(\mu_{a,f}, \sigma_{a,f}). \quad (2)$$

The initial belief state captures the knowledge the agent has about the general distribution of feature values in the environment. We assume that feature values are distributed $\text{Normal}(\mu_0, \sigma_0)$ and that the agent knows this. The initial belief state is thus defined $\mu_{a,f} = \mu_0$ and $\sigma_{a,f} = \sigma_0$ for all a and f .

Computations. A computational operation $c \in C$ corresponds to considering one feature of one option. Concretely, each computation measures the exact value of the feature and integrates that information into the belief state (as detailed in the next paragraph). We use $c_{a,f}$ to denote the computation that considers feature $x_{a,f}$. All meta-level MDPs additionally include a termination operation $\perp$, which denotes that computation should be terminated and an action should be selected based on the current belief state.

Meta-level transition function. The meta-level transition function T_{meta} describes how considering each feature updates the belief state. That is,

$$b_{t+1} \sim T_{\text{meta}}(b_t, c_t, s). \quad (3)$$

Each computation integrates the exact value of one feature into the belief state. If $c_t = c_{a,f}$, the updated belief b_{t+1} is identical to the previous belief b_t except that

$$\begin{aligned} \mu_{a,f} &= x_{a,f} \\ \sigma_{a,f} &= 0. \end{aligned} \quad (4)$$

Meta-level reward function. The reward function r_{meta} describes both the cost of computation and also the utility of the option that is ultimately chosen. For the former, we assume that considering different features of different options may have different costs, but that this cost does not depend on the belief or world state. Thus,

$$r_{\text{meta}}(b, c_{a,f}, s) = -\lambda_{a,f} \text{ for } c \neq \perp. \quad (5)$$

where $\lambda_{a,f}$ specifies the cost of considering feature f of option a . In our experiments, these costs will correspond to explicit information-gathering fees but the meta-level reward function can more generally capture the time and mental effort exerted while incorporating new information into one's beliefs.

The reward for terminating computation (i.e., executing $\perp$) is the reward associated with the external choice the agent makes based on their current belief state. It is defined as

$$r_{\text{meta}}(b, \perp, s) = r_{\text{object}}(s, a^*(b)) = \sum_f w_f x_{a^*(b),f} \quad (6)$$

where $a^*(b)$ is the action² the agent chooses; specifically, it is the action with maximal expected value given the current

belief state,

$$a^*(b) = \operatorname{argmax}_a \mathbb{E}_{s \sim b} [r_{\text{object}}(s, a)] \quad (7)$$

$$= \operatorname{argmax}_a \mathbb{E}_{(X, \mathbf{w}) \sim b} \left[\sum_f w_f x_{a,f} \right] \quad (8)$$

$$= \operatorname{argmax}_a \sum_f w_f \mu_{a,f}. \quad (9)$$

Thus, the meta-level reward for termination is the *true* utility of the action with maximal *estimated* utility.

Policy. Although not technically part of the meta-level MDP itself, in order to simulate cognitive processes and make behavioral predictions, we need one final component: the policy that selects which computation to execute in each possible belief state. One policy of special interest is the optimal policy, that is, the policy that maximizes meta-level return. However, meta-level MDPs typically have massive belief spaces that make computing the optimal policy intractable. To address this, early work in rational metareasoning proposed using a one-step lookahead approximation, termed the “meta-greedy” policy because it greedily maximizes meta-level reward (Russell & Wefald, 1991). Interestingly, the idea of approximating optimal selection of computations with a one-step lookahead was independently proposed by Gabaix, Laibson, Moloche, and Weinberg (2006) specifically in the context of multi-attribute choice. Here, we will also use this approximation. See Appendix B for a derivation of the meta-greedy policy for the meta-level MDP defined above.

Summary

In this section, we described a general formal framework for modeling decision-making as a sequential decision problem: meta-level MDPs. In this framework, a decision-making process is modeled as a sequence of basic information-processing operations (or “computations”) that an agent executes in order to update their beliefs about the values of different actions they can take (Figure 2B). Identifying an (approximately) optimal policy for a meta-level yields a resource-rational decision-making strategy—that is, a way to determine which information to consider and which to ignore on any given decision.

We then presented an application of the general framework to multi-attribute choice. In this simplified model, we assume that the utility of an option is a linear combination of its features, with the weights for each feature corresponding to the agent’s preferences. At the beginning of the decision, the agent does not know the values of the two options. Instead, they only have some sense of what values different features are likely to take have. This is formalized in their prior beliefs, which also form their initial *belief state*. Given this initial belief, they can only choose randomly among the

options. To make a better decision, the agent needs to deliberate. In particular, they can consider one of the feature values, and update their belief accordingly. We formalize this consideration as a *computational action* that moves the agent from their initial belief state to a new belief state, in which they are certain about the value of one feature and have a more precise (but still uncertain) belief about the utility of the corresponding option. Based on this new belief state, they must choose what to think about next (which computational action to execute), which will bring them to yet another belief state. However, each computation incurs a cost. Thus at some point, usually before considering all the possible information, the agent will *terminate computation*, choosing whichever option has the highest utility according to their final belief.

The power of our approach is that we can capture many different types of nudges with only minimal modifications to this model. By making the assumption that computations are selected rationally, the model can predict how small changes to a decision will influence the entire decision-making process. Nudges change which computations are best to execute, which will, in turn, affect the choices people make—sometimes dramatically so. In the remainder of the paper, we will apply the model to three different types of nudges. But first, we describe the paradigm we will use to simulate nudges in a controlled experimental environment.

An experimental paradigm for studying nudges

A key challenge for studying human decision-making (and by extension, the effect of choice architecture) is that the decision-making process is unobservable. To address this challenge, Payne, Bettman, and Johnson (1988) introduced the *Mouselab* paradigm, which makes participants’ decision-making processes observable. The basic idea is to occlude decision-relevant information and require participants to click on different areas of the screen to reveal it. Which pieces of information they uncover, and the order in which they do so, provides a highly detailed trace of their decision-making process.

In our version of the task (shown in Figure 4), participants are faced with a multi-attribute decision-making problem displayed as a table, with columns corresponding to choice options and rows corresponding to features on which the options vary. Concretely, the features correspond to different types of prizes, the values of which vary from trial to trial (shown in the leftmost column). To reveal the number of one type of prize for a given option, the participant must click the corresponding cell some number of times (the number may vary); we impose an explicit cost of one point per click.

²For notational clarity, we assume a single optimal action. In the actual model implementation, ties are broken randomly; thus, $a^*(b)$ is more precisely a uniform distribution over all optimal actions, and $r_{\text{meta}}(b, \perp, s)$ takes an expectation over them.

Figure 4*Experimental interface for Mouselab*

Prizes	Basket 1	Basket 2	Basket 3	Basket 4	Basket 5
A: 3 points	2		3	4	
B: 2 points	7				7
C: 2 points	7	4		2	
D: 21 points	7		8	6	
E: 2 points	9				6

Total click cost: 10 points

You won 2 A prizes, 7 B prizes, 7 C prizes, 7 D prizes, and 9 E prizes, totaling 199 points.

Total earnings (prize values minus click cost): \$0.063.

Note. On this problem, participants chose between five baskets, represented by the table columns. Each basket has five different prize types, with the value of each of these prizes given in the leftmost column. To reveal prize counts, participants could click on the corresponding red box once, with each click costing one point. At any point, participants could stop choosing and select a basket (e.g., basket 1 in this example). The participant then earned a bonus determined by the total value of the prizes in the selected basket minus the cost spent revealing boxes (30 points are worth one cent).

Besides making the decision-making process observable, the Mouselab paradigm has the convenient property of exactly externalizing the meta-level MDP for multi-attribute choice defined above. In particular, the table of prize numbers corresponds to the feature matrix X , the prize values correspond to the weights w (together, these form the state), the information currently visible in the table corresponds to a belief b , revealing a cell corresponds to a computation c , and the number of clicks necessary to reveal each cell corresponds to the meta-level reward function (specifically, $\lambda_{a,f}$ in Equation 5).

Using a paradigm that maps directly onto a meta-level MDP allows us to quantitatively evaluate the resource-rational approach to nudges without addressing the considerable challenge of modeling computational architectures for naturalistic decisions. Of course, addressing this challenge will be essential for the approach to be applied in practice, and it will be a critical direction for future research if our approach is found to be promising.

In the following sections, we illustrate how our approach can be used to understand three existing nudges: default options, suggesting alternatives, and information highlighting. We model each of these nudges as a modification to the basic meta-level MDP for multi-attribute choice described above.

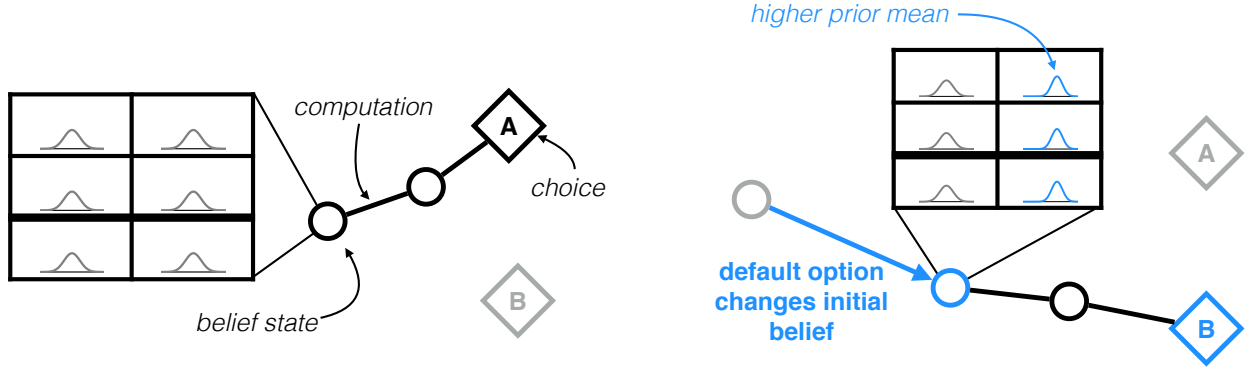
In each case, the modification has systematic consequences for the computations a resource-rational agent executes, and therefore on the choice they make. To test these predictions, we implement each class of nudge within the Mouselab paradigm, and compare the observed behavior with the simulations. Critically, all of the models we present are parameter-free, representing true *a priori* predictions of the resource-rational model. In each case, we confirm most or all of the key behavioral predictions.

Experiment 1: Default options

We begin with a simple but surprisingly effective class of nudges, default options. As the name suggests, these nudges involve changing which option people will get if they don't actively select another one. Although the efficacy of such nudges is due in part to changing the outcome for people that don't consider the decision at all, defaults still have an effect when people do deliberate (Dhingra, Gorn, Kener, & Dana, 2012). Here, we focus on this latter pathway.

Model

As illustrated in Figure 5, we model default nudges as modifications to the agent's initial belief state. This is consis-

Figure 5*Experiment 1: Formalizing default options*

Note. The decision maker assumes that the default option is the best choice for a typical person, so their initial belief assigns higher value to the default option’s features. This influences their choice both directly and indirectly—by changing which cognitive operations they execute. Circles represent belief states, lines represent cognitive operations, and diamonds represent choices. Hypothetical events are shown in gray and nudges are shown in light blue.

tent with the common assumption that the agent interprets the default as a recommendation (Johnson & Goldstein, 2003; Johnson et al., 2012; McKenzie et al., 2006). In particular, we assume that the default option is the one that is best for the “average” person, and further assume that the agent knows this and integrates the information provided by the default option accordingly. Intuitively, they will begin the decision-making process with the expectation that the features of the default option are better than average. In some cases, this initial expectation will lead the agent to avoid deliberation altogether and simply choose the default; in other cases, the agent will engage in some deliberation in order to tailor their choice to their own idiosyncratic preferences. However, even in the latter case, the default still influences their choice.

Recall that we model individual variability in preferences as different weights of a linear utility function (Equation 1). For simplicity, we assume that the weights are strictly positive and that each feature is equally important on average. The average preferences can then be characterized by $\mathbf{w} = \mathbf{1}$, and the default option for a decision problem with true feature values X is

$$d(X) = \operatorname{argmax}_a r_{\text{object}}((X, \mathbf{1}), a) = \operatorname{argmax}_a \sum_f x_{a,f}. \quad (10)$$

That is, the default option is the one with the greatest sum of unweighted feature values.

We assume that the agent knows this is how the default is chosen, and adjusts their initial belief accordingly. To maintain independent Gaussian beliefs about each feature value, we use a mean-field approximation to this belief update. That

is, the initial belief is updated with

$$(\mu_{a,f}, \sigma_{a,f}) = \begin{cases} (\mu^+, \sigma^+) & \text{if } a = d \\ (\mu^-, \sigma^-) & \text{if } a \neq d \end{cases}, \quad (11)$$

where μ^+ is the average value of a feature for an option that is best for the average person and μ^- is the same for an option that is *not* best. Likewise, σ^+ and σ^- are the standard deviation of the feature values in each case. We estimate these values numerically in one million simulated decision problems, separately for each problem size. As one would expect, this results in a higher prior mean for the features of the default and a lower prior mean for all other features. σ^+ and σ^- are both slightly lower than σ_0 .

To preview the results, the model predicts that people will be more likely to choose an option when it is presented as the default, and this effect is larger for more complex problems (those with more options and/or features). Importantly, this holds even in cases where the model does not immediately accept the default without performing any computation. Furthermore, the model predicts that providing a default option will never reduce the agent’s utility (including both choice payoff and computational cost), although it will be most beneficial to people who have typical preferences. We now test these predictions.

Methods

For this and all future experiments: data, code, pre-registrations, and experiment demos can be found at <https://github.com/fredcallaway/optimal-nudging>. All analyses were pre-registered, unless otherwise noted. All experiments were approved by the institutional review board of

Figure 6*Experiment 1: Example default-option trial*

Do you want to choose basket 3?

It pays the most when the prizes are equally valuable.

Yes
No

Prizes	Basket 1	Basket 2	Basket 3	Basket 4	Basket 5
A: 13 points					
B: 17 points					

Note. On a default trial (illustrated), participants could accept or reject a recommended basket that paid the most if the prizes were equally valuable (i.e., it identified the basket with the most prizes). If a participant did not accept the default, the banner was removed and participants could make their own choice. However, the basket label for the default option remained highlighted in green.

Princeton University (protocol number 10859), and all participants gave informed consent.

We tested our model predictions on Mouselab, with default options presented in a banner above the Mouselab table at the start of the trial (see Figure 6). The default option identified the basket that would pay the most if all the prizes were equally valuable—that is, the basket with the most prizes (ties were broken randomly). Participants were informed of this selection process and were reminded of it each time a default option as presented. When the default banner was presented, all click events for the Mouselab table were disabled and the default option was highlighted in green for visual saliency. Participants then chose between accepting the default immediately without revealing any prize counts, or making their own choice on the trial. If participants chose to make their own choice, the default banner was hidden for the remainder of the trial, but the basket label for the default option remained highlighted in green.

To determine prize counts, we sampled from a normal distribution with a mean of 5 and a standard deviation of 1.75, and then rounded and truncated these values so no counts were below 0 or above 10. To reveal prize counts, participants could click on the corresponding cell twice, incurring a cost of two points. Prize values were randomly sampled with the constraint that they sum to 30 points and each prize was worth at least one point.

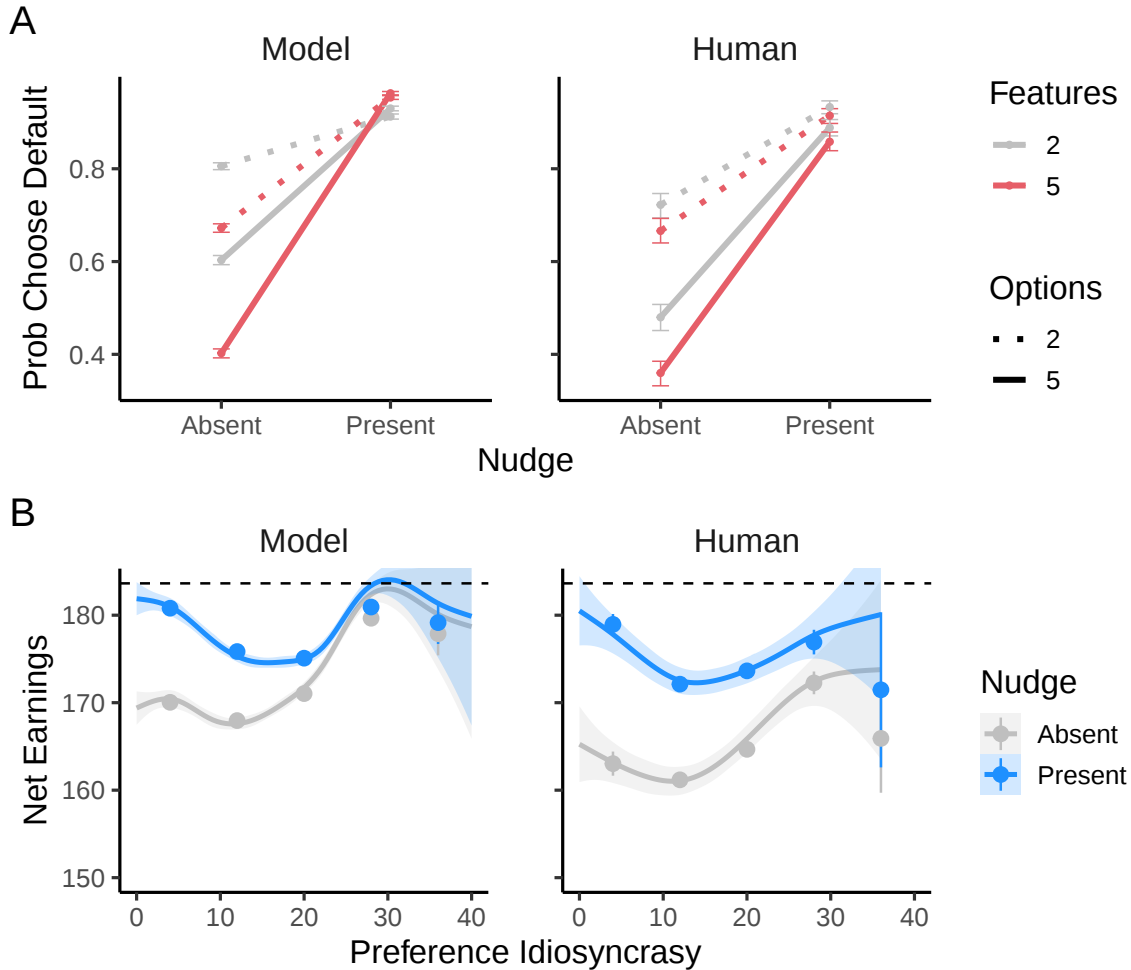
On each test trial, participants earned a bonus equal to the total value of the prizes in their selected basket minus the

points they spent revealing prize counts. At the end of the experiment, the total bonus each participant earned was paid to them as a bonus, with 30 points equaling one cent.

Participants completed 2 practice trials and 32 test trials. Half of the test trials were control trials, where no default option was presented, and half were nudge trials with a default option. Furthermore, half of the problems had two baskets (i.e., two Mouselab columns) and half had five baskets. Finally, half had two prize types (i.e., two Mouselab rows) and half had five prize types. There were thus three binary parameters determining each problem—trial type, number of baskets, and number of features. The stimuli were constructed so that each participant completed four problems for each of the eight unique parameter combinations, with trial ordering randomized. Participants earned \$1.30 for participating in the study plus an average bonus of \$1.79.

We recruited a preregistered sample size of 400 US-based participants from Prolific. This sample size was selected by performing a power analysis on simulated data from a meta-greedy decision maker on the same set of problems we used in the experiment. Participants who failed to pass a comprehension quiz in their first three tries were excluded from the experiment.

After collecting and analyzing the data, we discovered that some participants did not collect any information (and therefore chose randomly) on the majority of trials, even in the absence of a default option. While our pre-registered statistical tests were nonetheless significant, the effect sizes were sub-

Figure 7*Experiment 1: Model predictions and experimental results for default options*

Note. (A) The probability that the option which is best for the average person (the “default”) is chosen, depending on whether it is presented as the default option or not. Each line shows one problem complexity level, defined by the number of features and options. (B) Net earnings (payoff minus click cost) as a function of preference idiosyncrasy (L1 distance from the mean prize-value vector). The case with a default option is shown in blue. For this and all future results figures: the left panels show the prediction of the zero-parameter resource-rational model. The right panels show experimental data, excluding participants who revealed no information (and therefore chose a basket randomly) on more than half of control trials. Plots with full data are included in Appendix C. Points show binned means, error bars show 95% confidence intervals computed by bootstrapping, and regression lines show generative additive model fits with standard error confidence bands.

stantially reduced by the large amount of random responding. For this reason, we exclude all participants who made a choice without gathering any information on more than half of control trials (those without a default option). We apply the same exclusion criterion to all experiments, unless otherwise noted. Plots and statistics for the full dataset are provided in Appendix C. For Experiment 1, we excluded 102 participants (26%), leaving 298 participants in the analysis. We also excluded practice trials, as planned, leaving 9536

trials to conduct our pre-registered tests. As pre-registered, all reported p values reflect one-tailed tests to confirm the relevant model prediction.

Results

Figure 7A shows the probability that the option which is best for someone with average preferences is chosen, depending on whether it is presented as the default. In line with prior research, the model predicted—and our results

confirmed—that presenting an option as the default increases the probability of selecting it: participants chose the basket with the most prizes on 89.8% of trials when it was presented as the default option, compared with 55.7% on control trials. This difference was significant, as revealed in a logistic regression predicting default-chosen from nudge-present ($z = 34.77$, $p < .001$).

The model predicted that the default will be chosen without any deliberation on 74.7% of trials. However, the resource-rational agent is still more likely to choose the default after deliberating because the information encoded in the default option remains relevant as long as not all feature values have been considered. On the 25.3% of trials in which the model did not choose the default immediately, it chose the default on 76.2% of trials (compared to 62.1% for control trials). Consistent with this, in the 23.6% of trials on which our participant did not immediately choose the default, they were still more likely to select it eventually (63.7% vs. 57.8%; $z = 3.53$, $p < .001$). This suggests that participants' choices were affected by the informational content of the default over and above an automatic tendency to simply accept the default without any deliberation.

The model also predicted that the default would have a more pronounced effect for more complex decisions (operationalized as the number of options and features). Intuitively, this is because the default affects the agent's prior beliefs, and these priors play a stronger role when more feature values are left uncovered. However, the model predicted that people would be *less* likely to choose the default option when they had more idiosyncratic preferences (operationalized as the L1 distance of $\mathbf{w}$ from its mean value across individuals; recall that the preference weights, $\mathbf{w}$, are implemented in the experiment as the values of different types of prizes). To test these predictions, we ran a logistic regression predicting default-chosen with many-options and many-features (binary variables capturing whether there were five or two options or prizes, respectively), as well as idiosyncrasy (the L1 distance from the uniform weight vector), nudge-present, and the interaction of nudge-present with the other variables as independent variables. The interaction terms provide the critical tests, as they reveal the effect of the default after controlling for the main effects of problem complexity and idiosyncrasy on selection of the option with the most prizes. As predicted, we found significant positive interactions with many-options ($z = 5.06$, $p < .001$) and many-features ($z = 1.66$, $p = .049$), and a significant negative interaction with idiosyncrasy ($z = -5.87$, $p < .001$). Thus, consistent with the model's predictions, people were more likely to choose the default option (relative to baseline) when the problem was more complex, but they were less likely to choose it the more their preferences differed from average.

Finally, we investigated the effects of defaults on participant earnings. As illustrated in Figure 7B, the model pre-

dicted that defaults would be beneficial for everyone, but that this benefit would be largest for those with less idiosyncratic preferences (that is, trials with preferences weights closer to the uniform distribution). Consistent with this prediction, participants achieved higher net earnings (payoff minus click cost) when a default option was presented (174.44 points vs. 164.64 points; linear regression: $t(9534) = 14.60$, $p < .001$), but there was a significant negative interaction between nudge-present and idiosyncrasy ($t(9532) = -4.45$, $p < .001$).

Discussion

Our findings replicated and extended previous findings; presenting an option as the default increased the chance it was selected, and helped participants make better choices (Choi et al., 2006). Furthermore, default nudges were more effective on more complex choices: increasing the number of options and features increased the relative probability of selecting the default. However, default options were not effective for everybody—participants with more idiosyncratic preferences were less likely to choose the default, extending related findings that the impact of default options varies across groups (Beshears, Choi, Laibson, Madrian, & Wang, 2015; Löfgren et al., 2012). Finally, our model correctly predicted that people were more likely to choose the default not only when they made a choice without deliberation, but also on trials where they revealed feature values. In this way, our results unify the “cognitive effort” (Johnson & Goldstein, 2003; Johnson et al., 2012) and “recommendation” (Gigerenzer, 2008; McKenzie et al., 2006) theories of defaults in a common rational framework.

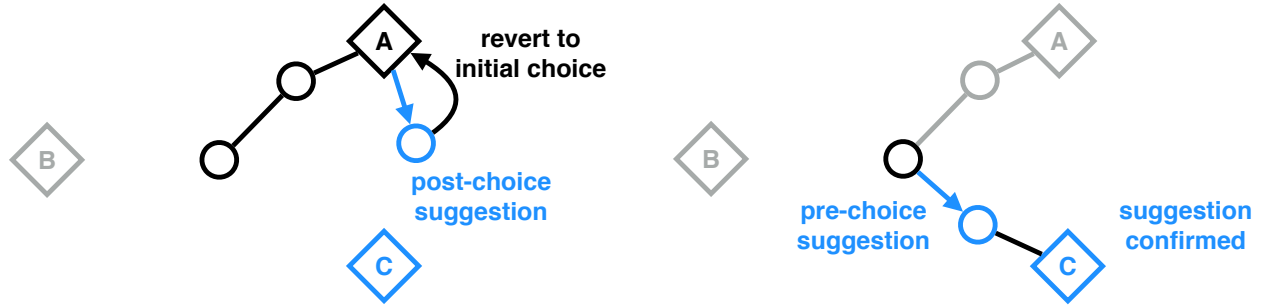
Experiment 2: Suggested alternatives

We now apply our framework to suggesting alternative options (ones that the agent would not have otherwise considered). Such suggestions could be made before the decision begins, as in recommender systems, or after an initial choice, as in up-sells. Here, we use our resource-rational framework to investigate and compare both settings.

Model

As illustrated in Figure 8, we model alternative suggestions as the addition of a new option, $\tilde{a}$, to the choice set. To capture the positive information that typically accompanies a suggestion (*try the kale; it's healthy!*), we assume that the best feature of the suggested option is immediately revealed to the agent at no cost.³ Formally, we identify the best feature

³One could also model the suggestion as a more general recommendation, as we did for defaults. We chose this alternative approach because suggested options are often not best for most people.

Figure 8*Experiment 2: Formalizing suggested alternatives*

Note. The suggestion is modeled as an addition of a new option to the choice set along with information about that option’s best feature. Suggestions made at the onset of the decision are more effective than those made after an initial choice (as is often done) because a late suggestion must override the previously considered information, while an early suggestion discourages considering other options at all. See Figure 5 for legend.

as

$$\tilde{f} = \underset{f}{\operatorname{argmax}} x_{\tilde{a},f}, \quad (12)$$

and we update the agent’s belief state by setting

$$\begin{aligned} \mu_{\tilde{a},f} &= x_{\tilde{a},f} \\ \sigma_{\tilde{a},f} &= 0. \end{aligned} \quad (13)$$

Note that this is equivalent to forcing the agent to execute $c_{\tilde{a},f}$.

We consider two versions of the suggestion nudge, which differ in when the suggestion is made. In the *early suggestion* version, we present the suggested option along with the original set, highlighting its best feature. In the *late suggestion* version, we first allow the agent to make an initial decision and then present the suggested option, giving them the option of changing their initial choice.

To preview the results, the model predicts that people will overall be more likely to choose the suggested option than chance, but that this suggestion will be more effective in complex environments. The model also predicts that early suggestions will be more effective than late suggestions. This is because early suggestions can influence the entire deliberation process.

Methods

We tested our predictions on alternative suggestions using a similar Mouselab setup as our defaults experiment. Suggestions were given either at the start of the trial (pre-choice suggestion) or after the participant has chosen a basket (post-choice suggestion). As in our defaults experiment, these suggestions were presented as a banner at the top of the screen. At the same time the banner was presented, the highest prize count for the suggested basket was revealed (ties were broken

randomly). Because suggested alternatives often involve introducing a new option, on trials with suggestions, there were six baskets to choose from, whereas there were only five on control trials (note that the sixth basket on post-choice suggestions was shown only after the participant has chosen a basket). On pre-choice trials, the suggested basket was chosen randomly from the six baskets. On post-choice trials, the right-most basket in the table was revealed and suggested. Unlike the defaults experiment, click events were not disabled when the banner was introduced, and the banner was displayed for the remainder of the trial after being shown.

Participants completed two practice trials and 30 test trials. No suggestion was given on control trials. Each problem had either two prize types or five prize types (i.e., features). The problems were arranged so that each participant completed 10 control problems (five with two features and five with five features), and 20 nudge trials (five for every unique combination of nudge timing and number of features), with problem order randomized. Participants earned \$1.30 for participating in the experiment plus an average bonus of \$1.66.

We recruited a preregistered sample size of 400 participants from Prolific, limiting our study to those living in the United States.⁴ As in Experiment 1, this sample size was selected by performing a power analysis on simulated data, and participants were required to pass a comprehension test in their first three tries. Prize counts, prize values, and bonuses were also determined in the same way as in Experiment 1.

We excluded 123 participants (31%) who gathered no in-

⁴We previously ran an experiment where some participants were told how we selected the revealed feature. However, participants had trouble understanding this process and so the present experiment does not include this information. Data for this experiment is available at <https://github.com/fredcallaway/optimal-nudging>.

Figure 9*Experiment 2: Example pre-choice suggestion trial*

Consider basket 1 - it has 6 B prizes!						
Prizes	Basket 1	Basket 2	Basket 3	Basket 4	Basket 5	Basket 6
A: 14 points						
B: 16 points	6					

Note. On pre-choice suggestion trials, a random basket at the start of the trial was chosen to be highlighted, and its highest feature value (i.e., prize count) was revealed. On post-choice suggestion trials, a new basket was revealed and highlighted using the same procedure after a participant chose a basket.

formation on more than half of control trials, leaving 277 participants in the analysis. We also excluded practice trials, as planned. We thus conducted our pre-registered tests on 8864 trials. As pre-registered, all reported p values reflect one-tailed tests to confirm the relevant model prediction.

Results

The key results are illustrated in Figure 10. Replicating previous findings and confirming our prediction, we found that participants chose suggested options significantly more often than chance, as measured by a chi-square test of independence (32.9% vs. 16.7%, $\chi^2(1) = 1052$, $p < .001$). The model predicts this effect because revealing the best feature of an option ensures that it is considered, while it might not have been otherwise.

As for default options, the model predicted that suggestions would be more effective for complex decisions. This prediction is somewhat counterintuitive because the single revealed feature has less weight when there are many features, suggesting that the suggestion should be *less* effective in this case. However, when there are few features, the model typically gathers all the information needed to choose an option with maximal or near-maximal value; thus there is little room for the suggestion to influence its choice. In the human data, suggestions were slightly more effective with five vs. two features ($z = 1.83$, $p = .034$), however the effect was small and it was not significant in the full (pre-exclusion) dataset ($z = 1.35$, $p = .089$).

The most striking model prediction concerns the relative efficacy of early and late suggestions. The model predicted that early suggestions would have a larger effect than late suggestions, especially for more complex problems. This effect occurs because the early suggestion can influence—or rather, preclude—later deliberation. When the suggestion is

made early, the agent may avoid the effort of searching for a better option, and simply takes the suggestion. In contrast, when the suggestion is made late, the agent will have already invested some effort into finding the option that is best on the features they value most. The positive information about the suggested item is unlikely to outweigh this earlier evidence. In line with our predictions, participants were more likely to choose the suggested option when it was presented before an initial choice ($z = -12.74$, $p < .001$). However, we did not observe a significant interaction with problem complexity ($z = -0.38$, $p = .353$).

Discussion

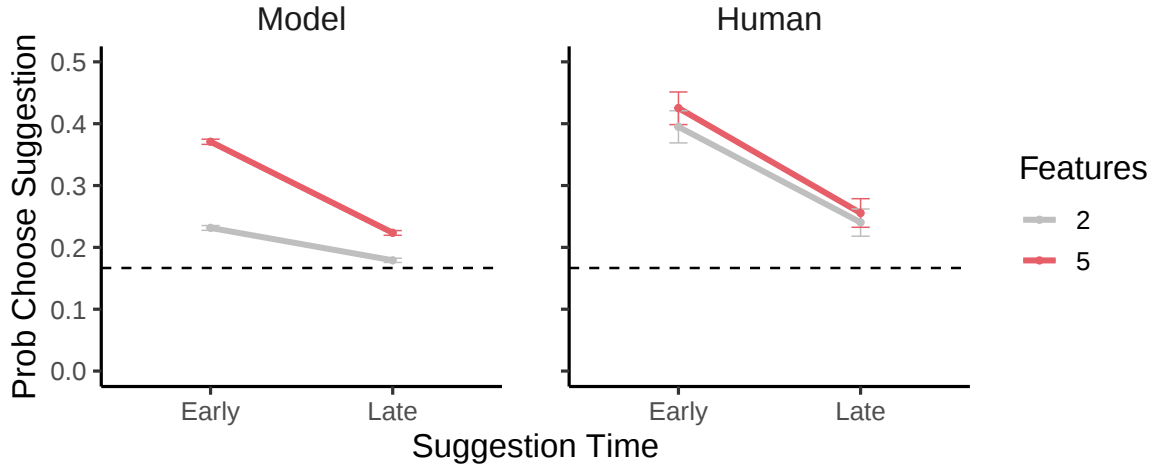
Replicating previous research, we found that suggesting an alternative option significantly increased the chance that participants would choose it. As we predicted, this effect was larger for suggestions that were given before a decision. This effect is largely driven by changing the deliberation process, or “script” for making a choice. This change in deliberation leads participants to perceive suggested options as unusually desirable, allowing us to capture two theories of suggested alternatives in a single model (Heidig et al., 2017; Schwartz et al., 2012). Our results may also shed light on why default options (Hummel & Maedche, 2019) and recommender systems (Häubl & Trifts, 2000) can be so effective—because they manipulate the choice architecture at the *start* of deliberation, they influence the entire decision process and can thus have substantial impacts on choice.

Experiment 3: Information highlighting

The final class of nudges we consider are those that draw the agent’s attention to specific features or options. This class of nudges extends the information-revealing mechanism we used to model suggestions to a case in which the nudge does

Figure 10

Experiment 2: Model predictions and experimental results for suggested alternatives



Note. Each panel shows the probability that the suggested option was chosen depending on whether it is presented at the beginning of the trial or after an initial decision has been made.

not force the agent to consider a specific piece of information, but instead makes it *easier* to consider some information.

Model

As illustrated in Figure 11, we model information highlighting as a reduction in the cost of certain computations. Intuitively, it is easier to consider information that is printed in large text on the front of a package versus small text on the back. Formally, this is captured in the $\lambda_{a,f}$ parameters (Equation 5), which assigns a cost to considering each feature of each option.

Many different types of nudges can be modeled as changes to the cost of certain computations. For example, to model a “foodscape” nudge, in which healthy foods are placed in prominent locations, we would reduce the computational cost of evaluating all the features for the healthy choices. Similarly, to model the “traffic light” nudge, in which the sugar and fat content of foods are prominently displayed, we would reduce the cost of evaluating the salt and fat feature for all the options. In our experiment, we consider a simplified form of the traffic light nudge in which only one feature is highlighted. Thus, on each trial we randomly select one feature, $\tilde{f}$, and set $\lambda_{a,\tilde{f}} = 1$ for all a , with all other values taking $\lambda_{a,f} = 3$.

The model predictions are straightforward. Because the agent knows the cost of each computation in advance, reducing the cost of a computation through information highlighting increases the chance that the agent will consider that information. As a result, the highlighted information will have

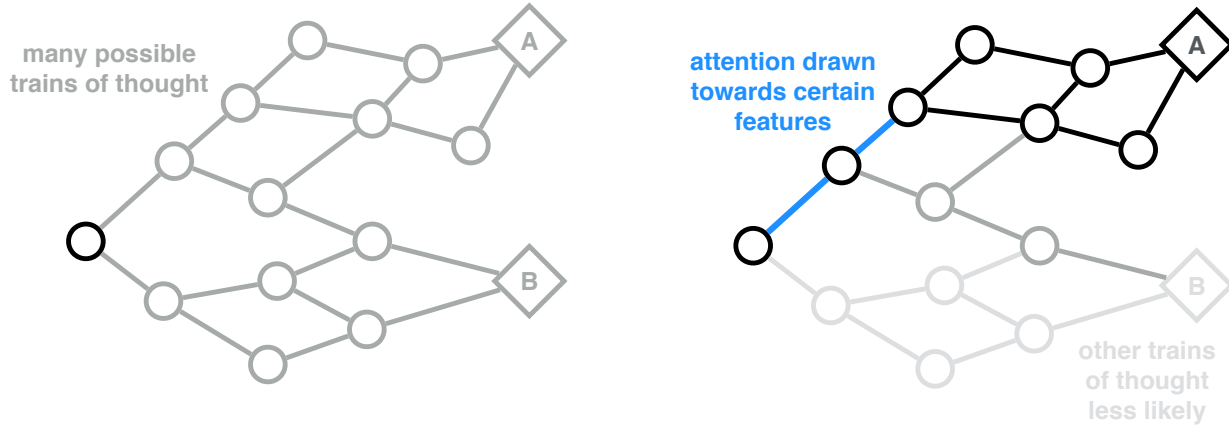
a greater impact on choice, and the agent will choose options that are better on the highlighted feature.

Methods

We tested our predictions on traffic light nudges using the same Mouselab process-tracing paradigm we used to study default nudges and suggestions. All problems had five baskets and three prize types, and all click costs were initially set to three points. Prize counts and bonuses were determined following the same procedure as the defaults experiment.

On every trial, one prize was randomly selected as the highlighted prize. On nudge trials, the click cost of all the highlighted prize’s values was reduced from three points to one point (see Figure 12). On control trials, the highlighted prize’s click costs were not changed. The value of the highlighted prize was sampled to achieve a close-to-uniform distribution for each participant. Concretely, on nudge trials, the value was sampled without replacement from either the set of even integers between 2 and 28 or the set of odd integers between 1 and 27 (both inclusive), with the set (even or odd) determined via simple randomization separately for each participant. The value of the highlighted prize on control trials was sampled without replacement from the complementary set of integers (i.e., if the highlighted prize values on nudge trials were sampled from the even integers, the highlighted prize values on control trials were sampled from the odd integers). On both control and nudge trials, the prize values of the two non-highlighted prizes were randomly sampled with the constraint that all three prize values sum to 30 points and each prize was worth at least one point.

Participants completed one practice nudge trial and one

Figure 11*Experiment 3: Formalizing information highlighting*

Note. When all information is equally accessible, many possible sequences of cognitive operations are possible and both choices are plausible. Highlighting some information reduces the cost of considering that information, leading the agent towards certain trains of thought and away from others. This in turn makes some choices more likely. See Figure 5 for legend.

Figure 12*Experiment 3: Example information-highlighting trial*

Clicks to reveal box		Prizes	Basket 1	Basket 2	Basket 3	Basket 4	Basket 5
1		A: 20 points					
2		B: 9 points					
3		C: 1 point					

Note. All click costs were initially set to three points. On nudge trials, the click costs for the highlighted prize's boxes were reduced to one point. On control trials, a highlighted prize was selected but its boxes were not put on sale.

practice control trial, then 14 test nudge trials and 14 test control trials. Trial order for both practice and test trials was randomized. Participants earned \$1.30 for completing the experiment, plus an average bonus of \$1.51.

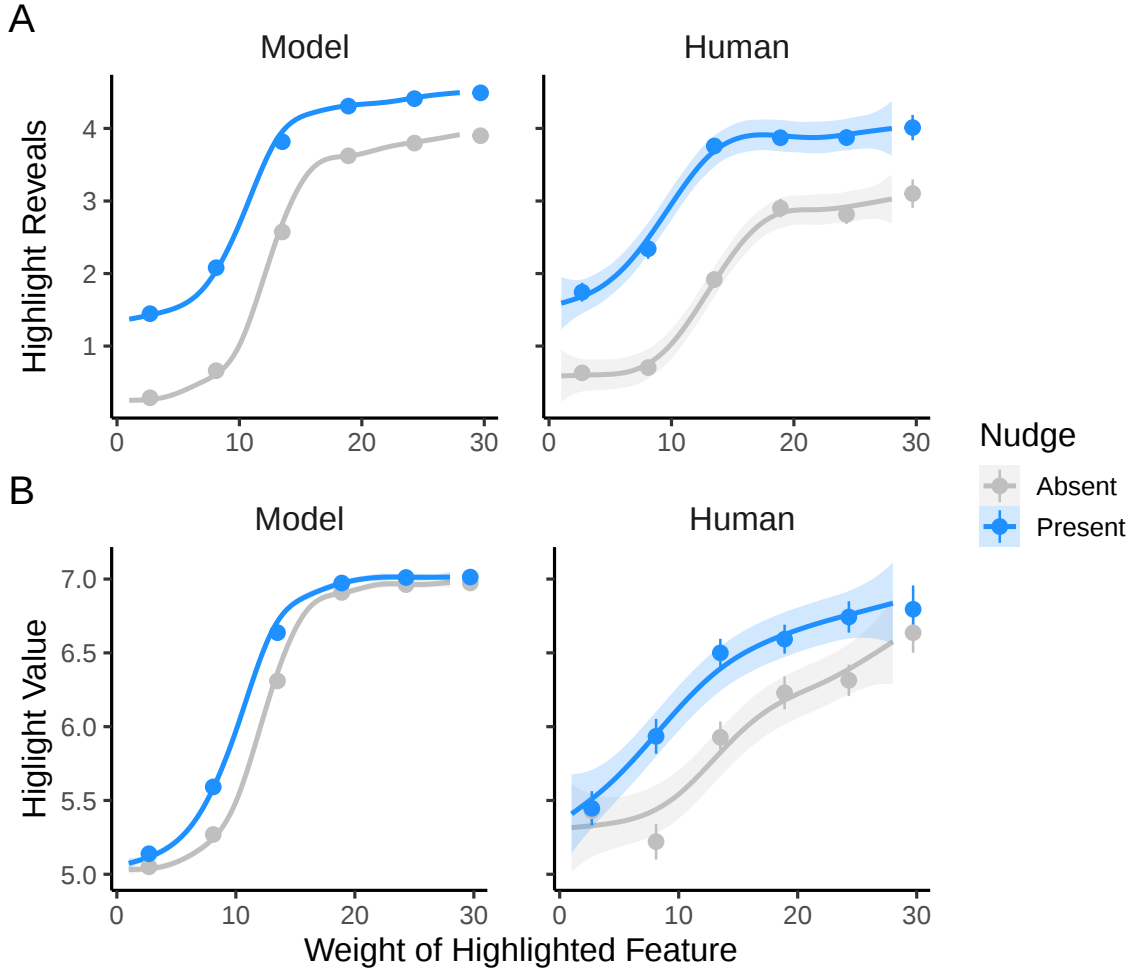
We recruited a preregistered sample size of 150 participants from Prolific, limiting our study to those living in the United States. As in the previous experiments, this sample size was selected by performing a power analysis on simulated data, and participants were required to pass a comprehension test in their first three tries. Prize counts, prize values, and bonuses were also determined in the same way as previous experiments.

We excluded 62 participants (41%) who gathered no in-

formation on more than half of control trials, leaving 88 participants in the analysis. We also excluded practice trials, as planned. We thus conducted our pre-registered tests on 2464 trials. As pre-registered, all reported p values reflect one-tailed tests to confirm the relevant model prediction.

Results

The model predicted that reducing the cost of considering a feature would increase the amount that people consider it. Indeed, as shown in Figure 13A, participants revealed an average of 3.20 values of the highlighted feature on nudge trials, compared with 1.89 values for control trials (two-sample

Figure 13*Experiment 3: Model predictions and experimental results for information highlighting*

Note. (A) The average number of cells that were revealed for the highlighted feature. (B) The average value of the highlighted feature for the chosen item. In both rows, the gray line shows a baseline value (for a randomly selected feature) when no feature is highlighted.

t-test: $t(2458.7) = 16.57, p < .001$).

In the model, revealing the value of a feature for more options effectively increases the weight of the feature; this leads it to choose options that have high value on that dimension. Figure 13B confirms this prediction. On nudge trials, participants chose baskets with an average of 6.29 prizes of the highlighted type, compared to a baseline of 5.89 on control trials ($t(2457.7) = 5.86, p < .001$). Similarly, participants chose the basket with the highest number of highlighted prizes significantly more often on nudge trials (67.4% vs. 54.4%; $\chi^2(1) = 44, p < .001$).

Discussion

Replicating earlier work, we found that reducing the cost of a feature had a large impact on participants' deliberation

strategies and choices (Sonnenberg et al., 2013). When the cost of a prize value was reduced to one point, participants revealed more values of that feature, chose options that had higher values for the highlighted feature, and were more likely to choose the option that maximized the highlighted feature. Crucially, we showed how each of these effects can result from a resource-rational strategy—"over-weighting" the highlighted option in one's choices can be optimal when individuals have limited time and attention.

While previous research has shown that nutritional labels are only effective for those who notice or use them (Ollberding, Wolf, & Contento, 2011), our results suggest a complementary causal structure—people who weight a feature (i.e., sugar content) highly are more likely to use information about that feature to make a decision. This means that while

labeling can reduce decision cost for those who value the highlighted feature, it may have a less pronounced impact on actual choices—highlighting a feature is most beneficial to those who would likely have already incorporated that feature into their decision. Those that do not weight the feature highly, by contrast, are less likely to utilize, or benefit from, the label. This suggests that studies using survey data—where people may report gains from reduced deliberation cost—may overestimate the effectiveness of labels relative to those that measure actual consumption (see Elbel, Kersh, Brescoll, and Dixon (2009); Elbel et al. (2013); Seward, Block, and Chatterjee (2016); Sonnenberg et al. (2013)).

Constructing optimal nudges

Until now, we have focused on using resource-rational analysis to understand and predict the effects of established nudges. Here, we go further, and use resource-rational analysis to design new nudges. Because the resource-rational model makes quantitative predictions about how a nudge will affect the agent’s decision-making process and eventual choice, we can define an objective function that takes as input a decision problem and a candidate nudge and returns a scalar indicating how desirable the agent’s behavior is expected to be under the modified choice architecture. We can then use an optimization algorithm to automatically identify the best possible nudge of a given class for a given problem.

The proposed method for constructing optimal nudges consists of five steps:

1. Model a decision problem as a meta-level MDP, M .
2. Specify a space of possible nudges as a set of possible modified meta-level MDPs, $\tilde{M}$.
3. Specify the goal of the nudge with an objective function g such that $g(\tilde{M}, s)$ specifies how desirable the decision maker’s behavior will be given the modified meta-level MDP $\tilde{M}$ if the true state of the world is s .
4. Specify the choice architect’s knowledge about the world as a distribution over possible states, b_{arch} .
5. Identify the optimal nudge as the modification that maximizes the expected value of the objective function, given the architect’s beliefs:

$$\tilde{M}^*(s) = \underset{\tilde{M} \in \tilde{\mathcal{M}}}{\operatorname{argmax}} \mathbb{E} \left[g(\tilde{M}, s) \mid s \sim b_{\text{arch}} \right] \quad (14)$$

Here, we illustrate this method in the context of multi-attribute choice.

Step 1: Metalevel MDP

The first step in our method for constructing optimal nudges is to model the decision-making process in the target domain as a meta-level MDP. We have already completed this step for multi-attribute choice as part of our analysis of existing nudges.

Step 2: Space of nudges

After modeling a target decision-making problem as a meta-level MDP, we next specify the space of possible nudges one could apply as a set of modified meta-level MDPs. We consider two classes of modifications—those that change the agent’s prior beliefs, and those that highlight information.

Changing the agent’s prior beliefs corresponds to modifying the initial belief state, b_0 . We have already seen two specific examples of this: we modeled the effect of defaults by increasing the prior mean for all features of the default option, and we modeled the effect of suggestions by immediately revealing the suggested option’s best feature (setting the prior to a delta distribution on that value). Here, we limit our attention to modifications of the latter type, i.e., those that correspond to immediately revealing the values of some features. However, rather than using a fixed rule (e.g., always revealing the best feature of a specific option), we instead allow for an arbitrary selection of feature values to reveal, with a constraint on the total number of revealed values. Concretely, we require that exactly three values are revealed. Each candidate nudge is identified by a set of three unique option-feature pairs $\{(a_1, f_1), (a_2, f_2), (a_3, f_3)\}$. $\tilde{M}$ is then identical to M except that $\mu_{a_i, f_i} = x_{a_i, f_i}$ and $\sigma_{a_i, f_i} = 0$ for $i \in \{1, 2, 3\}$.

Highlighting information corresponds to changing the meta-level reward function. We assume that only the costs can be modified, as the termination reward corresponds to the value of the chosen option. In the most general form, we could specify a unique reduction for every possible computation. However, this would result in a very large space of possible nudges. To create a more tractable space to optimize over, we apply the constraint that the cost is reduced by a fixed amount, δ , for exactly three feature values. Thus, as before, a nudge is identified by a set of three unique option-feature pairs $\{(a_1, f_1), (a_2, f_2), (a_3, f_3)\}$, modifying $\tilde{M}$ with $r_{\text{meta}}(b, c_{a_i, f_i}, s) = -(\lambda_{a, f} - \delta)$ for $i \in \{1, 2, 3\}$.

Step 3: Objective function

How should we select among the many possible nudges that we formalized in the previous step? Ideally, we would implement the nudge that best accomplishes our goals. But in order to specify which nudge best accomplishes our goals, we must first specify what exactly our goals are. In step 3, we make the goal of a nudge mathematically precise in the form of an objective function.

There are many types of goals a nudge might have. For example, many nudges aim to maximize the probability that people take a certain action, e.g., recycling or registering to become an organ donor. This kind of goal can be formalized as maximizing the probability of the decision maker choos-

ing a specific action,

$$g_{\text{action}}(\tilde{M}, s; a) = \mathbb{E}_{b_T} [a^*(b_T) = a \mid \tilde{M}, s], \quad (15)$$

where the expectation is taken with respect to the belief state of the decision maker when they make a choice, b_T . Although the distribution over b_T depends on $\tilde{M}$ and s in complex ways, reducing the cost of a computation will generally make the agent more likely to execute it. Thus, this objective function will select modifications that make it inexpensive to consider features that make the desired action, a , appear desirable or make competing actions seem undesirable.

Other nudges, such as suggestions from digital recommendation systems, aim not to make people choose a specific option, but rather to improve the overall quality of their decisions. We can model this kind of goal as maximizing the expected utility of the decision maker's choice,

$$f_{\text{utility}}(\tilde{M}, s) = \mathbb{E}_{b_T} [\text{utility}(s, a^*(b_T)) \mid \tilde{M}, s]. \quad (16)$$

Finally, a choice architect might want to not only encourage people to make *better* decisions, but also to make it *easier* to make those decisions. We can formalize this goal as maximizing the cumulative meta-level reward, which captures both decision quality and computational cost,

$$f_{\text{meta}}(\tilde{M}, s) = \mathbb{E}_{\mathbf{c}} \left[\sum_t^T r(b_t, c_t) \mid \tilde{M}, s \right]. \quad (17)$$

Here, the expectation is taken over all possible sequences of computations the agent could execute. This quantity is also called the *meta-level return*, and it is the quantity that the optimal meta-level policy maximizes. We chose this objective for the experiments presented below.

Step 4: Architect belief

As formalized in the previous step, the desirability of a nudge depends on the true state of the world. The choice architect will typically know something about that state, ideally information that individual agents don't have direct access to (for example, the average annual out-of-pocket costs for a given insurance policy). At the same time, the agent may have access to information that the architect lacks (for example, their own risk preferences). We thus specify the choice architect's knowledge of the world as a distribution over world states, b_{arch} , that may differ from the agent's initial belief.

For our experiments, we will assume that the architect has perfect knowledge of the feature values, but does not know the agent's preference weights. We do assume, however, that the architect knows the distribution from which $\mathbf{w}$ is drawn. In our experiments, this is a uniform distribution over all possible integer weights that sum to thirty. We thus marginalize

over this distribution when computing the expected value of a nudge in Equation 14. Because there are a huge number of possible weights, we approximate the expectation with 1000 Monte Carlo samples.

Step 5: Optimization

Given a model of a decision-making process (step 1), a space of nudges (step 2), an objective function specifying the goal of the nudge (step 3), and a belief about the true state of the world (step 4), the final step is to identify the nudge that maximizes the objective function. When the number of possible nudges is relatively small, this can be achieved by exhaustive enumeration. However, to take full advantage of the flexibility of our approach, we must be able to specify very large spaces of nudges. To show that the method is robust to the use of imperfect optimization algorithms, we employ a simple hill climbing procedure. Recall that both spaces of nudges we consider correspond to selecting a set of three cells in the payoff matrix (to either immediately reveal or to reduce the cost of measuring). We begin by considering all the nudges in which only a single cell is included in the set, selecting the cell that maximizes the objective function (breaking ties randomly). We then commit to including this cell in the set, and repeat the process with the next cell, choosing the one that results in the best performance when added to the set with the first selected cell. Finally, we repeat the process once more to select the third cell. We emphasize that this procedure is not guaranteed to find the truly optimal nudge; however, this will only weaken our results. Developing better tools to optimize over less-constrained spaces of nudges is an important direction for future work.

Experiment 4: Optimal nudging by modifying beliefs

We first apply our optimal nudging procedure to develop nudges that directly modify a decision maker's initial belief state. In Mouselab, this corresponds to revealing a set of feature values at the beginning of a trial. To demonstrate the value of our approach over and above simply making information more accessible, we compare optimal nudges against two baselines: a weak baseline in which features are revealed randomly, and a strong baseline in which the most extreme feature values are revealed (for a similar procedure, see Cioffi, Levitsky, Pacanowski, & Bertz, 2015).

Methods

We tested the efficacy of optimal belief-modifying nudges in Mouselab, but with some prize values revealed immediately. All problems had five options and five features, with prize values generated following the same method as our other experiments. On each problem, click costs were initially set to two. Before applying any nudge, we first selected three random prize counts (i.e., table cells) to reveal. For

Figure 14*Experiment 4: Problem construction procedure***1.) Generate payoff matrix and set all but three random costs to two points**

Basket 1	Basket 2	Basket 3	Basket 4	Basket 5	Basket 1	Basket 2	Basket 3	Basket 4	Basket 5
4	5	3	7	2			3		
7	3	7	6	7					7
6	7	4	5	6					
7	7	5	3	6					
5	3	3	6	7					7

**2.) Generate optimal, extreme, and random modifications****Optimal modifications**

Basket 1	Basket 2	Basket 3	Basket 4	Basket 5
		3		2
				7
6				
7				
				7

Extreme modifications

Basket 1	Basket 2	Basket 3	Basket 4	Basket 5
		3	7	2
				7
	7			
				7

Random modifications

Basket 1	Basket 2	Basket 3	Basket 4	Basket 5
		3		
				7
7	7			
			6	7

**3.) On each trial, select modification type and sample feature weights**

Prizes	Basket 1	Basket 2	Basket 3	Basket 4	Basket 5
A: 8 points			3		2
B: 1 point					7
C: 11 points	6				
D: 5 points	7				
E: 5 points					7

Note. The weights were not known when constructing the nudges.

each nudge type, three additional prize counts were revealed. These values were chosen randomly for random nudges, and for extreme nudges, the three hidden values furthest from 5 were revealed (ties were broken randomly). For optimal nudges, the three values were selected to maximize expected meta-level total reward (bonus minus click cost), integrating over possible prize values (see Figure 14). Before the experiment, we generated 1,000 Mouselab problems, and constructed optimal, extreme, and random modifications for each. To generate stimuli, we first sampled from the generated problems without replacement, and then selected a nudge (optimal, extreme, or random) to determine the initially revealed prize counts.

Participants completed two practice trials and 30 test trials. Participants made decisions on 10 test trials for each modification type, with trial order randomized. Practice trials were generated following a similar procedure, but always had random modifications. Participants earned \$1.30 for participating in the study plus an average bonus of \$1.66.

We recruited a preregistered sample size of 250 US-based participants from Prolific. As in Experiment 1, this sample size was selected by performing a power analysis on simulated data, and participants were required to pass a comprehension test in their first three tries.

Because all trials had some values revealed immediately, we did not exclude any of the 250 participants. However, we did exclude practice trials, as planned. We thus conducted our pre-registered tests on 7500 trials. As pre-registered, all reported p values reflect one-tailed tests to confirm the relevant model prediction.

Results

On average, participants earned 161.3 points on trials with random nudges, 166.1 points on trials with the heuristic nudge that revealed the most extreme values, and 169.5 points on trials with optimal nudges. A linear regression with optimal nudge trials as the reference group revealed that performance was significantly worse in the other two groups (random: $t(7497) = -10.55$, $p < .001$; heuristic: $t(7497) = -4.42$, $p < .001$). This overall performance benefit of optimal nudges was supported by a significant increase in decision quality (value of the chosen basket) and a significant decrease in decision cost (clicking penalty). On average, participants chose baskets worth 172.8 points on trials with optimal nudges, compared to 169.6 points with heuristic nudges ($t(7497) = -4.13$, $p < .001$) and 165.0 points with random nudges ($t(7497) = -10.03$, $p < .001$). At the same time, they incurred a clicking penalty of 3.3 points on trials with optimal nudges, compared to 3.5 points with heuristic nudges ($t(7497) = 1.69$, $p = .045$) and 3.7 points with random nudges ($t(7497) = 3.05$, $p = .001$).

Discussion

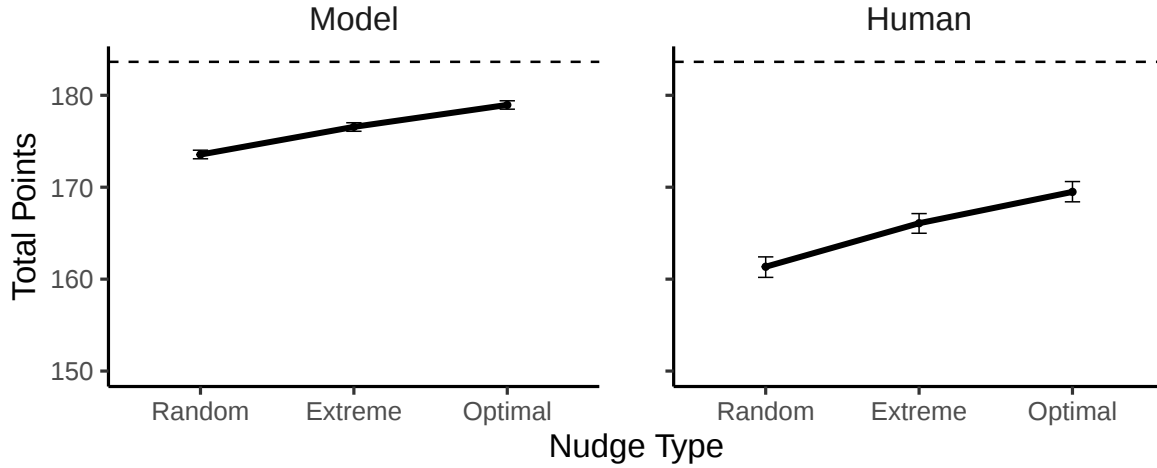
Our findings replicate previous work showing that not all information highlighting nudges are equally effective at improving choice (Lin et al., 2017). Indeed, we found that compared to extreme and random modifications, nudges determined by our procedure increased participants' total reward and improved the quality of their choices. Furthermore, optimal nudging made these choices easier to make—compared to trials with extreme and random modifications, participants spent significantly fewer points revealing prize counts on trials with optimal nudges.

Our approach to constructing information highlighting nudges has a number of additional advantages over other approaches. First, we explicitly specify the goal of the nudge using an objective function. This can increase the transparency of the nudge, provide a natural way to think about new types of goals for nudges (e.g., making people's decisions easier without systematically changing their choices), and allow individuals to have control over when, how, and why they are nudged. Second, given a model of the decision-making process and an objective, our method automatically discovers an optimal nudge using computational optimization techniques. This method has the potential to improve information highlighting by identifying novel choice architectures, but can also reduce the human labor and cost involved in designing these nudges. Along these lines, we believe that our optimal nudging procedure may be especially useful for constructing choice architectures in digital environments (Weinmann, Schneider, & Vom Brocke, 2016). Finally, by integrating over all possible preferences, our model can be applied in heterogeneous populations or in domains where people's preferences are unknown or unstable (Payne, Bettman, & Johnson, 1992; Slovic, 1995). This can potentially address criticisms that nudges constructed for the "average" decision maker can be ineffective or even harmful for certain subgroups (Costa & Kahn, 2013; Peer et al., 2020; Thunström, Gilbert, & Ritten, 2018).

Experiment 5: Optimal nudging by modifying costs

In many domains, it may be infeasible to directly manipulate people's belief states. For example, in high-information environments, individuals may focus on only one or two features of a choice (Kalnikaitė et al., 2013), making it difficult for choice architects to deterministically manipulate attention. Indeed, food labeling interventions that aim to modify people's beliefs states by providing additional information are often unsuccessful (Lin et al., 2017). Instead, effective information highlighting nudges generally reduce the cost of evaluating key pieces of information. By increasing the chance that certain features are considered, these nudges have an indirect and stochastic influence on people's beliefs.

Here, we test our optimal nudging framework on infor-

Figure 15*Experiment 4: Model predictions and experimental results for optimal belief modification*

Note. Each point shows the average number of points attained under different strategies for generating belief-modification nudges. The dashed line shows the average maximal option value (i.e., the utility achieved by an unboundedly rational agent). The expected value of choosing an option randomly is 150.

mation highlighting nudges that reduce costs without modifying the initial belief state. In the Mouselab setup, this corresponds to reducing the cost of certain prize counts at the start of a trial, rather than fully revealing these values. As in the previous experiment, we will compare optimal cost-reductions to both random and extremity-based cost-reductions.

Methods

All aspects of the design were identical to Experiment 4 with two exceptions. Click costs were initially set to three points (rather than two points), and the cost for some cells were reduced to one point (rather than being revealed entirely).

We recruited a preregistered sample size of 250 US-based participants from Prolific. Participants earned \$1.30 for participating in the study plus an average bonus of \$1.58. To maintain comparability with Experiment 4, we did not exclude any of our 250 participants in our analyses. We did exclude practice trials, as planned. We thus conducted our pre-registered tests on 7500 trials. As pre-registered, all reported p values reflect one-tailed tests to confirm the relevant model prediction.

Results

As in the previous experiment, we found that our procedure helped participants choose more valuable baskets with fewer clicks, resulting in higher total payoffs. On average, participants earned 162.1 points on trials with opti-

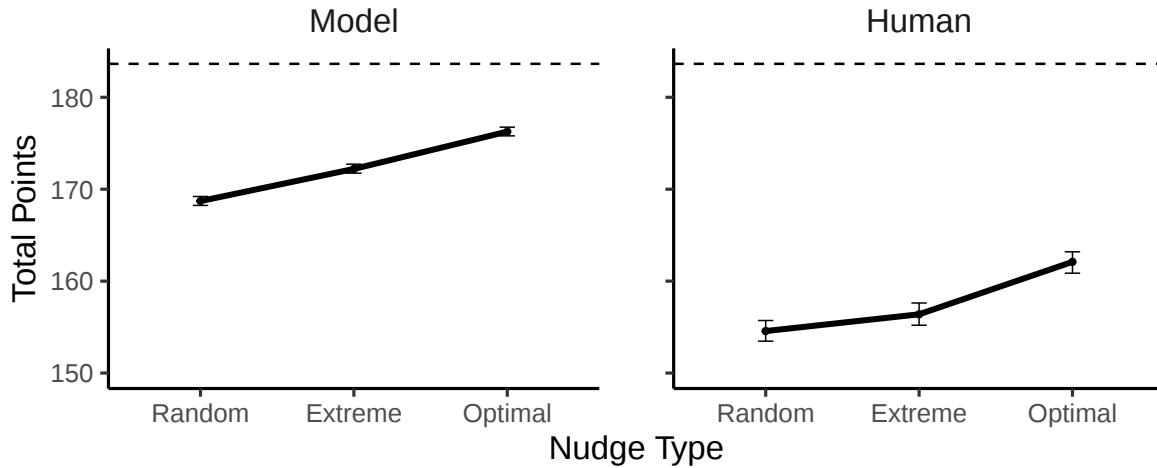
mal nudges, compared to 156.4 points with heuristic nudges ($t(7497) = -6.73$, $p < .001$) and 154.6 points with random nudges ($t(7497) = -8.87$, $p < .001$). They chose baskets worth 166.3 points on trials with optimal nudges, compared to 161.1 points with heuristic nudges ($t(7497) = -6.22$, $p < .001$) and 159.4 points with random nudges ($t(7497) = -8.18$, $p < .001$). Finally, they incurred a clicking penalty of 4.2 points on trials with optimal nudges, compared to 4.7 points with heuristic nudges ($t(7497) = 2.35$, $p = .009$) and 4.8 points with random nudges ($t(7497) = 3.20$, $p < .001$).

Discussion

Similar to our findings on belief-state nudging, we found that our optimal nudging procedure significantly improved participants' choices. Compared to random and extreme modifications, participants chose better baskets and spent fewer points revealing prize counts on trials with optimal modifications. Crucially, these nudges were effective even though they influenced participants' beliefs only indirectly. This setup likely better reflects modern real-world environments where individuals are often distracted, hurried, and inundated with information (Roetzel, 2019). Our results thus highlight the flexibility of our framework—optimal nudges can be constructed for any problem where the choice architect can specify a detailed model of deliberation and identify a suitable space of possible nudges.

Figure 16

Experiment 5: Model predictions and experimental results for optimal cost reduction



Note. Each point shows the average number of points attained under different strategies for generating cost-reduction nudges. The dashed line shows the average maximal option value (i.e., the utility achieved by an unboundedly rational agent). The expected value of choosing an option randomly is 150.

General Discussion

In this paper, we proposed a formal framework for modeling, constructing, and evaluating nudges. Our approach is based on theoretical work that characterizes human decision-making as making optimal use of limited computational resources (Griffiths et al., 2015; Lieder & Griffiths, 2020). In this framework, nudges change the initial belief state or sequence of deliberative actions taken by a resource-rational decision-maker. This in turn influences their observed choices. While models have traditionally been developed separately for different nudges and contexts (Chetty, 2015; Yeung, 2012), we are able to account for the effects of several nudges in this framework. Furthermore, because our model has no free parameters, we are able to make our predictions before observing any human behavior.

We tested our framework in five large behavioral experiments. In the first three experiments, we showed how resource-rational analysis can be used to make precise predictions about the effects of three commonly-used nudges: default options, suggested alternatives, and information highlighting. In each case, our model both identified novel phenomena and replicated findings from applied research, allowing us to unify several verbal theories of specific nudges in a common formal framework.

We then showed how our resource-rational approach to nudging can be used to automatically construct *optimal* nudges. Our approach identified nudges that were significantly more effective than those identified randomly or by a heuristic. While we chose to optimize meta-level reward,

or overall well-being, this approach could easily be extended to optimize other types of nudges, such as those that maximize the probability of making a certain choice or those that reduce deliberation cost without systematically changing behavior.

Furthermore, because our optimal nudging procedure is automatic, our approach could be used to extend the concept of *personalized* nudges (Mills, 2020; Peer et al., 2020; Schöning, Matt, & Hess, 2019; Sunstein, 2013, 2014; Thaler & Tucker, 2013; Yeung, 2017). Traditional approaches to personalizing nudging use past choices or other user data to estimate people’s preferences, circumstances, and needs. In our framework, one could infer a user’s preference within a single decision based on observable measures of their decision-making operations (for example, mouse- or eye-tracking). On the other hand, when data privacy is a concern (Mills, 2020), our approach can still be applied without any user data by integrating over all possible preferences as we did in our experiments. Similarly, our approach could automate the construction of *self-nudges* (Reijula & Hertwig, 2020), or choice architectures that individuals manipulate and design to help improve their own decisions. In our framework, self-nudges could be constructed by allowing individuals to specify the objective of the nudge or even the space of possible modifications. This transparency and individual autonomy could potentially address arguments that nudging can be manipulative and paternalistic (Goodwin, 2012; Hausman & Welch, 2010; Wilkinson, 2013).

Despite the advantages, our framework for modeling nudging presents several challenges that should be addressed

in future work. First, it requires a detailed model of the computations underlying the decision we would like to intervene on. In the present work, we avoided this challenge by using a process-tracing paradigm that externalized these typically unobservable processes. Applying the method in the real world, however, requires one to infer this model from behavior. Nevertheless, even a heavily simplified decision-making model may be adequate to make useful, if not highly accurate, predictions. At the very least, our framework identifies novel predictions and choice architectures nudges that researchers can test in applied domains.

Second, the effect of a nudge can often be formalized in the model in numerous ways, and the best formalization may vary by domain. For example, while our model of defaults—assuming that they provide information about which option is best for most people—may work well in the context of choosing healthcare plans, it may not work when individuals and the choice architect have highly divergent goals. Thus, even with our approach, models may still have to be adjusted for different domains and contexts, limiting the potential for full automation. Nevertheless, our approach still greatly constrains the space of possible models, and provides a general framework for designing and comparing different possible models of a given nudge.

Finally, applying the method requires identifying an (approximately) optimal decision-making strategy given the assumed cognitive architecture. Following previous work in computer science (Hay et al., 2012; Matheson, 1968; Russell & Wefald, 1991), economics (Gabaix et al., 2006; Hébert & Woodford, 2017), neuroscience (Drugowitsch et al., 2012; Jang, Sharma, & Drugowitsch, 2021; Tajima et al., 2016) and psychology (Callaway, Lieder, et al., 2018; Callaway et al., 2021; Chen et al., 2021; Howes et al., 2016), our solution method assumes that decision-making can be modeled as the generation of information that is optimally incorporated into a belief state. However, many models of human decision-making cannot easily be cast in this way, as they involve arbitrary transformations and operations on the decision-relevant information (see Bhatia, Loomes, & Read, 2021 for a recent review). Fortunately, the assumption of Bayesian beliefs is not a requirement for the general approach. As long as a decision-making model can be specified as a sequential process in which cognitive operations update mental states, near-optimal cognitive processes can be identified by model-free reinforcement learning (Callaway, Gul, Krueger, Griffiths, & Lieder, 2018).

Nudges have already proven to be a highly effective mechanism for improving the decisions people make. In proposing a formal framework for modeling the effects of choice architecture, we hope to provide insight into how we can design even more effective nudges. In addition to providing a computational tool for predicting the effects of nudges, this framework forces us to confront important questions about

what the goals of nudging are. Together, these advances allow us to apply tools from artificial intelligence to automate the design of nudges. We anticipate that this will make it possible to increase the range of contexts in which nudges can be used. More broadly, our approach of understanding nudges as modifications to a decision-maker's internal computational environment may have implications for the larger goal of designing interfaces that support human decision-making.

Acknowledgements

This research was supported by NSF grant number 1930720.

References

- Abadie, A., & Gay, S. (2006). The impact of presumed consent legislation on cadaveric organ donation: A cross-country study. *Journal of health economics*, 25(4), 599–620.
- Allcott, H., & Kessler, J. B. (2019). The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics*, 11(1), 236–76.
- Antinyan, A., & Asatryan, Z. (2019). *Nudging for tax compliance: A meta-analysis* (Discussion Paper No. 19-055). ZEW-Centre for European Economic Research.
- Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., ... Galing, S. (2017). Should governments invest more in nudging? *Psychological Science*, 28(8), 1041–1055.
- Bergeron, S., Doyon, M., Saulais, L., & Labrecque, J. (2019). Using insights from behavioral economics to nudge individuals towards healthier choices when eating out: A restaurant experiment. *Food Quality and Preference*, 73, 56–64.
- Berkowitsch, N. A., Scheibehenne, B., & Rieskamp, J. (2014). Rigorously testing multialternative decision field theory against random utility models. *Journal of Experimental Psychology: General*, 143(3), 1331–1348.
- Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., & Wang, S. Y. (2015). *Who is easier to nudge?* (Vol. 401; NBER Working Paper). National Bureau of Economic Research.
- Bhatia, S., Loomes, G., & Read, D. (2021). Establishing the laws of preferential choice behavior. *Judgment and Decision Making*, 16(6), 46.
- Bhatia, S., & Stewart, N. (2018). Naturalistic multiattribute choice. *Cognition*, 179, 71–88.
- Bhui, R., Lai, L., & Gershman, S. J. (2021, October). Resource-rational decision making. *Current Opinion in Behavioral Sciences*, 41, 15–21.
- Bothos, E., Apostolou, D., & Mentzas, G. (2015). Recommender systems for nudging commuters towards eco-friendly decisions. *Intelligent Decision Technologies*, 9(3), 295–306.
- Botvinick, M. M., Niv, Y., & Barto, A. G. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, 113(3), 262–280.
- Botvinick, M. M., & Toussaint, M. (2012). Planning as inference. *Trends in Cognitive Sciences*, 16(10), 485–488.

- Brown, J. R., Farrell, A. M., & Weisbenner, S. J. (2012). *The downside of defaults* (NBER Working Paper). National Bureau of Economic Research.
- Bucher, T., Collins, C., Rollo, M. E., McCaffrey, T. A., De Vlieger, N., Van der Bend, D., . . . Perez-Cueto, F. J. (2016). Nudging consumers towards healthier choices: A systematic review of positional influences on food choice. *British Journal of Nutrition*, 115(12), 2252–2263.
- Buckley, F. H. (2009). *Fair governance: Paternalism and perfectionism*. Oxford University Press.
- Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends in Cognitive Sciences*, 23(3), 251–263.
- Cai, C. W. (2020). Nudging the financial market? A review of the nudge theory. *Accounting & Finance*, 60(4), 3341–3365.
- Callaway, F., Gul, S., Krueger, P., Griffiths, T. L., & Lieder, F. (2018). Learning to select computations. In *Uncertainty in Artificial Intelligence: Proceedings of the Thirty-Fourth Conference*.
- Callaway, F., Lieder, F., Das, P., Gul, S., Krueger, P. M., & Griffiths, T. L. (2018). A resource-rational analysis of human planning. In *Proceedings of the 40th annual conference of the cognitive science society*.
- Callaway, F., Rangel, A., & Griffiths, T. L. (2021). Fixation patterns in simple choice reflect optimal information sampling. *PLOS Computational Biology*, 17(3), e1008863.
- Carlsson, F., & Johansson-Stenman, O. (2019). *Optimal prosocial nudging* (Working Paper No. 757). University of Gothenburg.
- Carroll, G. D., Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2009). Optimal defaults and active decisions. *The Quarterly Journal of Economics*, 124(4), 1639–1674.
- Chen, H., Chang, H. J., & Howes, A. (2021, May). Apparently Irrational Choice as Optimal Sequential Decision Making. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(1), 792–800.
- Chetty, R. (2015). Behavioral economics and public policy: A pragmatic perspective. *American Economic Review*, 105(5), 1–33.
- Choi, J., Laibson, D., Madrian, B., Metrick, A., McCaffrey, E., & Slemrod, J. (2006). Saving for retirement on the path of least resistance. In *Behavioral public finance: Toward a new agenda* (p. 304–351). New York: Russell Sage Foundation.
- Cioffi, C. E., Levitsky, D. A., Pacanowski, C. R., & Bertz, F. (2015). A nudge in a healthy direction. The effect of nutrition labels on food purchasing behaviors in university dining facilities. *Appetite*, 92, 7–14.
- Cohen, A. L., Kang, N., & Leise, T. L. (2017). Multi-attribute, multi-alternative models of choice: Choice, reaction time, and process tracing. *Cognitive Psychology*, 98, 45–72.
- 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(3), 680–702.
- Damgaard, M. T., & Nielsen, H. S. (2018). Nudging in education. *Economics of Education Review*, 64, 313–342.
- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12), 1704–1711.
- Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective and Behavioral Neuroscience*, 8(4), 429–453.
- Dayan, P., & Huys, Q. J. M. (2008). Serotonin, inhibition, and negative mood. *PLOS Computational Biology*, 4(2), e4.
- Deliza, R., & MacFie, H. (2001). Product packaging and branding. In L. J. Frewer, E. Risvik, & H. Schifferstein (Eds.), *Food, people and society* (pp. 55–72). Springer.
- Dhingra, N., Gorn, Z., Kener, A., & Dana, J. (2012). The default pull: An experimental demonstration of subtle default effects on preferences. *Judgment & Decision Making*, 7(1), 69–76.
- 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(4), 332–341.
- Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience*, 32(11), 3612–3628.
- Elbel, B., Kersh, R., Brescoll, V. L., & Dixon, L. B. (2009). Calorie labeling and food choices: A first look at the effects on low-income people in New York City. *Health Affairs*, 28(Supplement 1), w1110–w1121.
- Elbel, B., Mijanovich, T., Dixon, L. B., Abrams, C., Weitzman, B., Kersh, R., . . . Ogedegbe, G. (2013). Calorie labeling, fast food purchasing and restaurant visits. *Obesity*, 21(11), 2172–2179.
- Ellison, B., Lusk, J. L., & Davis, D. (2014). The impact of restaurant calorie labels on food choice: Results from a field experiment. *Economic Inquiry*, 52(2), 666–681.
- Elsweiler, D., Trattner, C., & Harvey, M. (2017). Exploiting food choice biases for healthier recipe recommendation. In *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval* (pp. 575–584).
- Farhi, E., & Gabaix, X. (2020). Optimal taxation with behavioral agents. *American Economic Review*, 110(1), 298–336.
- Felsen, G., & Reiner, P. B. (2015). What can neuroscience contribute to the debate over nudging? *Review of Philosophy and Psychology*, 6(3), 469–479.
- Forget, A., Chiasson, S., Van Oorschot, P. C., & Biddle, R. (2008). Improving text passwords through persuasion. In *Proceedings of the 4th symposium on usable privacy and security* (pp. 1–12).
- Forwood, S. E., Ahern, A. L., Marteau, T. M., & Jebb, S. A. (2015). Offering within-category food swaps to reduce energy density of food purchases: A study using an experimental online supermarket. *International Journal of Behavioral Nutrition and Physical Activity*, 12(1), 1–10.
- Freire, W. B., Waters, W. F., & Rivas-Mariño, G. (2017). Nutritional traffic light system for processed foods: Qualitative study of awareness, understanding, attitudes, and practices in Ecuador. *Revista Peruana de Medicina Experimental y Salud Pública*, 34(1), 11–18.
- Freire, W. B., Waters, W. F., Rivas-Mariño, G., Nguyen, T., & Ri-

- vas, P. (2017). A qualitative study of consumer perceptions and use of traffic light food labelling in Ecuador. *Public Health Nutrition*, 20(5), 805–813.
- Fryer Jr, R. G., Levitt, S. D., List, J., & Sadoff, S. (2012). *Enhancing the efficacy of teacher incentives through loss aversion: A field experiment* (Tech. Rep.). National Bureau of Economic Research.
- Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96(4), 1043–1068.
- Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245), 273–278.
- Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, 3(1), 20–29.
- Goda, G. S., Levy, M. R., Manchester, C. F., Sojourner, A., & Tasoff, J. (2015). *The role of time preferences and exponential-growth bias in retirement savings* (Tech. Rep.). National Bureau of Economic Research.
- Goodwin, T. (2012). Why we should reject “nudge”. *Politics*, 32(2), 85–92.
- 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(5), 829–846.
- Gottlieb, J., Oudeyer, P. Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, 17(11), 585–593.
- Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, 7(2), 217–229.
- Hamerman, E. J., Rudell, F., & Martins, C. M. (2018). Factors that predict taking restaurant leftovers: Strategies for reducing food waste. *Journal of Consumer Behaviour*, 17(1), 94–104.
- Häubl, G., & Trifts, V. (2000). Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Science*, 19(1), 4–21.
- Hausman, D. M., & Welch, B. (2010). Debate: To nudge or not to nudge. *Journal of Political Philosophy*, 18(1), 123–136.
- Hay, N., Russell, S., Tolpin, D., & Shimony, S. E. (2012). Selecting computations: Theory and applications. In *Proceedings of the 28th conference on uncertainty in artificial intelligence*.
- Hébert, B., & Woodford, M. (2017). Rational inattention with sequential information sampling. *Working Paper*, 1–141.
- Heidig, W., Wentzel, D., Tomczak, T., Wiecek, A., & Falzl, M. (2017). “Supersize me!” The effects of cognitive effort and goal frame on the persuasiveness of upsell offers. *Journal of Service Management*, 28(3), 541–562.
- Howes, A., Lewis, R. L., & Vera, A. (2009). Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review*, 116(4), 717–751.
- Howes, A., Warren, P. A., Farmer, G., El-Dereby, W., & Lewis, R. L. (2016). Why contextual preference reversals maximize expected value. *Psychological Review*, 123(4), 368–391.
- Huh, Y. E., Vosgerau, J., & Morewedge, C. K. (2014). Social defaults: Observed choices become choice defaults. *Journal of Consumer Research*, 41(3), 746–760.
- Hummel, D., & Maedche, A. (2019). How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics*, 80, 47–58.
- Hunt, L. T., Rutledge, R. B., Malalasekera, W. M. N., Kennerley, S. W., & Dolan, R. J. (2016). Approach-induced biases in human information sampling. *PLoS Biology*, 14(11), e2000638.
- Huys, Q. J. M., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., ... Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, 112(10), 3098–3103.
- Jachimowicz, J. M., Duncan, S., Weber, E. U., & Johnson, E. J. (2019). When and why defaults influence decisions: A meta-analysis of default effects. *Behavioural Public Policy*, 3(2), 159–186.
- Jang, A. I., Sharma, R., & Drugowitsch, J. (2021, March). Optimal policy for attention-modulated decisions explains human fixation behavior. *eLife*, 10, e63436.
- Jesse, M., & Jannach, D. (2021). Digital nudging with recommender systems: Survey and future directions. *Computers in Human Behavior Reports*, 3, 100052.
- Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, 302(5649), 1338–1339.
- Johnson, E. J., & Goldstein, D. G. (2004). Defaults and donation decisions. *Transplantation*, 78(12), 1713–1716.
- Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7(1), 35–51.
- Johnson, E. J., Shu, S. B., Dellaert, B. G. C., Fox, C., Goldstein, D. G., Häubl, G., ... Weber, E. U. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, 23(2), 487–504.
- Jung, J. Y., & Mellers, B. A. (2016). American attitudes toward nudges. *Judgment & Decision Making*, 11(1), 62–74.
- Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1-2), 99–134.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.
- Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics*, 112(2), 375–406.
- Kalnikaite, V., Bird, J., & Rogers, Y. (2013). Decision-making in the aisles: Informing, overwhelming or nudging supermarket shoppers? *Personal and Ubiquitous Computing*, 17(6), 1247–1259.
- Keramati, M., Dezfouli, A., & Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology*, 7(5), e1002055–e1002055.
- Kiszkó, K. M., Martinez, O. D., Abrams, C., & Elbel, B. (2014). The influence of calorie labeling on food orders and con-

- sumption: A review of the literature. *Journal of community health*, 39(6), 1248–1269.
- Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological Science*, 28(9), 1321–1333.
- Kosters, M., & Van der Heijden, J. (2015). From mechanism to virtue: Evaluating nudge theory. *Evaluation*, 21(3), 276–291.
- Krukowski, R. A., Harvey-Berino, J., Kolodinsky, J., Narsana, R. T., & DeSisto, T. P. (2006). Consumers may not use or understand calorie labeling in restaurants. *Journal of the American Dietetic Association*, 106(6), 917–920.
- Lehner, M., Mont, O., & Heiskanen, E. (2016). Nudging—A promising tool for sustainable consumption behaviour? *Journal of Cleaner Production*, 134, 166–177.
- Lewis, R. L., Howes, A., & Singh, S. (2014). Computational rationality: Linking mechanism and behavior through bounded utility maximization. *Topics in Cognitive Science*, 6(2), 279–311.
- Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43.
- Lieder, F., Krueger, P. M., & Griffiths, T. (2017). An automatic method for discovering rational heuristics for risky choice. In *Proceedings of the 39th annual meeting of the cognitive science society*.
- Lin, Y., Osman, M., & Ashcroft, R. (2017). Nudge: Concept, effectiveness, and ethics. *Basic and Applied Social Psychology*, 39(6), 293–306.
- Liu, P. J., Wisdom, J., Roberto, C. A., Liu, L. J., & Ubel, P. A. (2014). Using behavioral economics to design more effective food policies to address obesity. *Applied Economic Perspectives and Policy*, 36(1), 6–24.
- Loewenstein, G., Asch, D. A., Friedman, J. Y., Melichar, L. A., & Volpp, K. G. (2012). Can behavioural economics make us healthier? *BMJ*, 344, e3482.
- Löfgren, Å., Martinsson, P., Hennlock, M., & Sterner, T. (2012). Are experienced people affected by a pre-set default option—Results from a field experiment. *Environmental Economics and Management*, 63(1), 66–72.
- Löfgren, Å., & Nordblom, K. (2020). A theoretical framework of decision making explaining the mechanisms of nudging. *Journal of Economic Behavior & Organization*, 174, 1–12.
- Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *The Quarterly Journal of Economics*, 116(4), 1149–1187.
- Matheson, J. E. (1968). The economic value of analysis and computation. *IEEE Transactions on Systems Science and Cybernetics*, 4(3), 325–332.
- McKenzie, C. R., Liersch, M. J., & Finkelstein, S. R. (2006). Recommendations implicit in policy defaults. *Psychological Science*, 17(5), 414–420.
- Mills, S. (2020). Personalized nudging. *Behavioural Public Policy*, 1–10.
- Momsen, K., & Stoerk, T. (2014). From intention to action: Can nudges help consumers to choose renewable energy? *Energy Policy*, 74, 376–382.
- Moseley, A., & Stoker, G. (2013). Nudging citizens? Prospects and pitfalls confronting a new heuristic. *Resources, Conservation and Recycling*, 79, 4–10.
- Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53(3), 139–154.
- Noguchi, T., & Stewart, N. (2018). Multialternative decision by sampling: A model of decision making constrained by process data. *Psychological Review*, 125(4), 512–544.
- O'Donoghue, T., & Rabin, M. (1998). Procrastination in preparing for retirement. , 125–156.
- Ollberding, N. J., Wolf, R. L., & Contento, I. (2011). Food label use and its relation to dietary intake among us adults. *Journal of the American Dietetic association*, 111(5), S47–S51.
- Orozco, F., Ochoa, D., Muquínche, M., Padro, M., & Melby, C. L. (2017). Awareness, comprehension, and use of newly-mandated nutrition labels among mestiza and indigenous Ecuadorian women in the central Andes region of Ecuador. *Food and Nutrition Bulletin*, 38(1), 37–48.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 534–552.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology*, 43(1), 87–131.
- Peer, E., Egelman, S., Harbach, M., Malkin, N., Mathur, A., & Frik, A. (2020). Nudge me right: Personalizing online security nudges to people's decision-making styles. *Computers in Human Behavior*, 109, 106347.
- Reijula, S., & Hertwig, R. (2020). Self-nudging and the citizen choice architect. *Behavioural Public Policy*, 1–31.
- Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, 108(2), 370–392.
- Roetzl, P. G. (2019). Information overload in the information age: A review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development. *Business Research*, 12(2), 479–522.
- Ronayne, D., & Brown, G. D. A. (2017). Multi-attribute decision by sampling: An account of the attraction, compromise and similarity effects. *Journal of Mathematical Psychology*, 81, 11–27.
- Rostain, T. (2000). Educating homo economicus: Cautionary notes on the new behavioral law and economics movement. *Law & Society Review*, 34(4), 973–1006.
- Russell, S., & Wefald, E. (1991). Principles of metareasoning. *Artificial Intelligence*, 49(1-3), 361–395.
- Russo, J. E., & Doshier, B. A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 9(4), 676–696.
- Sandoval, L. A., Carpio, C. E., & Sanchez-Plata, M. (2019). The effect of “traffic-light” nutritional labelling in carbonated soft drink purchases in Ecuador. *PloS one*, 14(10), e0222866.
- Schafer, J. B., Konstan, J., & Riedl, J. (1999). Recommender systems in e-commerce. In *Proceedings of the 1st ACM conference on electronic commerce* (pp. 158–166).

- Schöning, C., Matt, C., & Hess, T. (2019). Personalised nudging for more data disclosure? On the adaption of data usage policies format to cognitive styles. In *Proceedings of the 52nd Hawaii international conference on system sciences*.
- Schwartz, J., Riis, J., Elbel, B., & Ariely, D. (2012). Inviting consumers to downsize fast-food portions significantly reduces calorie consumption. *Health Affairs*, 31(2), 399–407.
- Seward, M. W., Block, J. P., & Chatterjee, A. (2016). A traffic-light label intervention and dietary choices in college cafeterias. *American Journal of Public Health*, 106(10), 1808–1814.
- Shu, S. B., & Gneezy, A. (2010). Procrastination of enjoyable experiences. *Journal of Marketing Research*, 47(5), 933–944.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118.
- Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3), 665–690.
- Sinclair, S. E., Cooper, M., & Mansfield, E. D. (2014). The influence of menu labeling on calories selected or consumed: A systematic review and meta-analysis. *Journal of the Academy of Nutrition and Dietetics*, 114(9), 1375–1388.
- Slovic, P. (1995). The construction of preference. *American Psychologist*, 50(5), 364–371.
- Solway, A., Diuk, C., Córdova, N., Yee, D., Barto, A. G., Niv, Y., & Botvinick, M. M. (2014). Optimal behavioral hierarchy. *PLOS Computational Biology*, 10(8), 1–10.
- Sonnenberg, L., Gelsomin, E., Levy, D. E., Riis, J., Barraclough, S., & Thorndike, A. N. (2013). A traffic light food labeling intervention increases consumer awareness of health and healthy choices at the point-of-purchase. *Preventive Medicine*, 57(4), 253–257.
- Starke, A. D., Kløverød Brynstad, E. K., Hauge, S., & Løke-land, L. S. (2021). Nudging healthy choices in food search through list re-ranking. In *Adjunct proceedings of the 29th ACM conference on user modeling, adaptation and personalization* (pp. 293–298).
- Steel, P. (2007). The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological bulletin*, 133(1), 65.
- Sunstein, C. R. (2013). *Impersonal default rules vs. active choices vs. personalized default rules: A triptych* (SSRN Scholarly Paper No. ID 2171343). Social Science Research Network.
- Sunstein, C. R. (2014). *Why nudge? The politics of libertarian paternalism*. Yale University Press.
- Sunstein, C. R. (2016). The council of psychological advisers. *Annual Review of Psychology*, 67(1), 713–737.
- Sunstein, C. R. (2017). Nudges that fail. *Behavioural public policy*, 1(1), 4–25.
- Sunstein, C. R. (2019). Nudging: A very short guide. *Business Economics*, 54(2), 127–129.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
- Szaszi, B., Palinkas, A., Palfi, B., Szollosi, A., & Aczel, B. (2018). A systematic scoping review of the choice architecture movement: Toward understanding when and why nudges work. *Journal of Behavioral Decision Making*, 31(3), 355–366.
- Tajima, S., Drugowitsch, J., & Pouget, A. (2016). Optimal policy for value-based decision-making. *Nature Communications*, 7(1), 12400.
- Tannenbaum, D., Fox, C. R., & Rogers, T. (2017). On the misplaced politics of behavioural policy interventions. *Nature Human Behaviour*, 1(7), 1–7.
- Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow: Using behavioral economics to increase employee saving. *Journal of Political Economy*, 112(S1), S164–S187.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- Thaler, R. H., & Tucker, W. (2013). Smarter information, smarter consumers. *Harvard Business Review*, 91(1), 44–54.
- Thorndike, A. N., Riis, J., Sonnenberg, L. M., & Levy, D. E. (2014). Traffic-light labels and choice architecture: Promoting healthy food choices. *American journal of Preventive Medicine*, 46(2), 143–149.
- Thunström, L., Gilbert, B., & Ritten, C. J. (2018). Nudges that hurt those already hurting—distributional and unintended effects of salience nudges. *Journal of Economic Behavior & Organization*, 153, 267–282.
- Tomov, M. S., Schulz, E., & Gershman, S. J. (2021). Multi-task reinforcement learning in humans. *Nature Human Behaviour*, 5(6), 764–773.
- Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multi-attribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, 121(2), 179–205.
- Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review*, 111(3), 757–769.
- van Kleef, E., van den Broek, O., & van Trijp, H. C. (2015). Exploiting the spur of the moment to enhance healthy consumption: Verbal prompting to increase fruit choices in a self-service restaurant. *Applied Psychology: Health and Well-Being*, 7(2), 149–166.
- Vercellis, C. (2009). *Business intelligence: Data mining and optimization for decision making*. Wiley Online Library.
- Vlaev, I., King, D., Dolan, P., & Darzi, A. (2016). The theory and practice of “nudging”: Changing health behaviors. *Public Administration Review*, 76(4), 550–561.
- Voyer, B. (2015). ‘Nudging’ behaviours in healthcare: Insights from behavioural economics. *British Journal of Healthcare Management*, 21(3), 130–135.
- Weinmann, M., Schneider, C., & Vom Brocke, J. (2016). Digital nudging. *Business & Information Systems Engineering*, 58(6), 433–436.
- White, B., Jiang, D., & Albarracin, D. (2021). The limits of defaults: The influence of decision time on default effects. *Social Cognition*, 39(5), 543–569.
- Wilkinson, T. M. (2013). Nudging and manipulation. *Political Studies*, 61(2), 341–355.
- Willis, L. E. (2013). When nudges fail: Slippery defaults. *The University of Chicago Law Review*, 80, 1155–1229.
- Wilson, A. L., Buckley, E., Buckley, J. D., & Bogomolova, S. (2016). Nudging healthier food and beverage choices through salience and priming: evidence from a systematic review. *Food Quality and Preference*, 51, 47–64.

- Xiao, B., & Benbasat, I. (2007). E-commerce product recommendation agents: Use, characteristics, and impact. *MIS Quarterly*, 31(1), 137–209.
- Yeung, K. (2012). Nudge as fudge. *The Modern Law Review*, 75(1), 122–148.
- Yeung, K. (2017). ‘Hypermudge’: Big data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118–136.

Appendix A Metalevel MDPs

Metalevel MDPs extend the standard MDP formalism to model the sequential decision problem posed by resource-bounded computation. We define a meta-level MDP as $(\mathcal{S}, \mathcal{A}, r_{\text{object}}, \mathcal{B}, C, T_{\text{meta}}, r_{\text{meta}})$. The first three components define the task-level problem. They have the same interpretation as $\mathcal{S}$, $\mathcal{A}$ and r in a standard MDP (we omit the transition function because we limit our analysis to one-shot decisions). The latter four components define the meta-level problem. We now define these four components in turn.

Beliefs. A belief state $b \in \mathcal{B}$ captures the agent’s current knowledge about the relevant state of the world. Formally, a belief is a distribution states, $\mathcal{B} \subseteq \Delta(\mathcal{S})$. Note that $\Delta(\mathcal{S})$ denotes the set of all possible distributions over $\mathcal{S}$. Importantly, contrary to a standard rational treatment of beliefs, the belief states in a meta-level MDP do not include all the information that is available to the DM. Instead, the belief state only contains information that is immediately accessible, excluding, for example, long-term memories and the number of calories in every box of cereal on a shelf.

Computations. A computational operation $c \in C$ is a primitive operation afforded by the computational architecture. Formally, it is a meta-level action that updates the belief in much the same way as a regular action changes state. All meta-level MDPs include the termination operation $\perp$, which denotes that computation should be terminated and an action should be selected based on the current belief state. We further explain belief updating and termination in the following two paragraphs.

Transition function. The meta-level transition function $T_{\text{meta}} : \mathcal{B} \times C \times \mathcal{S} \rightarrow \Delta(\mathcal{B})$ describes how computation updates beliefs. At each time step, the next belief is sampled from a distribution that depends on the current belief, the computational operation that was just executed, and the true state of the world, that is,

$$b_{t+1} \sim T_{\text{meta}}(b_t, c_t, s). \quad (\text{A1})$$

The transition function thus defines the core structure of the computational architecture. Following previous work (Hay et al., 2012; Matheson, 1968), we assume that the effect of computation is to generate or reveal information about the true state of the world, which is then integrated into the be-

lief state. Thus, in expectation, computation has the effect of making one’s beliefs more precise and accurate, although an individual computation may yield misleading information.

Reward function. The meta-level reward function $r_{\text{meta}} : \mathcal{B} \times C \times \mathcal{S} \rightarrow \mathbb{R}$ describes both the costs and benefits of computation. For the former, r_{meta} assigns a strictly negative reward for all non-terminating computational operations,

$$r_{\text{meta}}(b, c, s) = -\text{cost}(c) \text{ for } c \neq \perp. \quad (\text{A2})$$

The cost of computation may include multiple factors, such as energetic costs and opportunity costs.

Intuitively, the benefit of computation is that it allows one to make better decisions. This is captured by the meta-level reward for the termination operation $\perp$, defined as the true utility of the external action that the DM would execute given the current belief. We assume that the action is selected optimally. Thus,

$$r_{\text{meta}}(b, \perp, s) = r_{\text{object}}(s, a^*(b)). \quad (\text{A3})$$

where

$$a^*(b) = \underset{a}{\operatorname{argmax}} \mathbb{E} \left[r_{\text{object}}(s, a) \mid s \sim b \right] \quad (\text{A4})$$

In English, the meta-level reward for termination is the *true* utility of the action⁵ with maximal *estimated* utility.

Policy. The solution to a meta-level MDP takes the form of a policy $\pi : \mathcal{B} \rightarrow \Delta(C)$ that (perhaps stochastically) selects which computation to perform in each possible belief state. The optimal policy is the one that maximizes expected meta-level return,

$$\pi^* = \underset{\pi}{\operatorname{argmax}} \mathbb{E} \left[\sum_{t=1}^T r_{\text{meta}}(B_t, C_t, S) \mid C_t \sim \pi \right]. \quad (\text{A5})$$

Unfortunately, computing an exact optimal policy is intractable for problems of even moderate complexity. However, a greedy approximation to the optimal policy can achieve reasonable performance, sufficient for a predictive model of human behavior. We show how this greedy approximation can be derived in the next appendix.

Appendix B Meta-greedy policy

The basic intuition behind the meta-greedy policy of Russell and Wefald (1991) is to use one-step look-ahead in a transformed *belief MDP* (Kaelbling, Littman, & Cassandra, 1998), where the true state is marginalized out.

⁵For notational clarity, we have assumed a single optimal action. When multiple actions have the same expected value, we assume that ties are broken randomly; thus, $a^*(b)$ is more precisely a uniform distribution over all optimal actions, and $r_{\text{meta}}(b, \perp, s)$ takes an expectation over them.

Belief MDP

Given a meta-level MDP, $(\mathcal{S}, \mathcal{A}, r_{\text{object}}, \mathcal{B}, C, T_{\text{meta}}, r_{\text{meta}})$, we can derive a new MDP in which the first three components are integrated into the transition and reward functions. The result is a standard MDP where states are beliefs and actions are computations. To accomplish this, we must derive versions of T_{meta} and r_{meta} that marginalize over the true state of the world.

Marginal reward function. The marginal reward function is defined

$$r_{\text{meta}}(b, c) = \mathbb{E}_{s \sim b} [r_{\text{meta}}(b, s, c)]. \quad (\text{B1})$$

For $c \neq \perp$, $r_{\text{meta}}(b, s, c)$ does not depend on s , and we have simply

$$r_{\text{meta}}(b, c) = -\text{cost}(c). \quad (\text{B2})$$

The reward for terminating, however, depends on the state of the world; we must marginalize it out. Replacing $r_{\text{meta}}(b, s, \perp)$ with its definition, we have

$$r_{\text{meta}}(b, \perp) = \mathbb{E}_{s \sim b} [r_{\text{object}}(s, a^*(b))] \quad (\text{B3})$$

$$= \mathbb{E}_{s \sim b} \left[r_{\text{object}}(s, \arg\max_a \mathbb{E}_{s' \sim b} [r_{\text{object}}(s', a)]) \right] \quad (\text{B4})$$

$$= \max_a \mathbb{E}_{s \sim b} [r_{\text{object}}(s, a)]. \quad (\text{B5})$$

(B5) follows from

$$f(\arg\max_x f(x)) = \max_x f(x), \quad (\text{B6})$$

where f is $\mathbb{E}_{s \sim b} [r_{\text{object}}(s, \cdot)]$. To derive the specific expression for the multi-attribute model, we replace r_{object} with its definition, giving us

$$r_{\text{meta}}(b, \perp) = \max_a \mathbb{E}_{(X, \mathbf{w}) \sim b} \left[\sum_f w_f x_{a,f} \right] \quad (\text{B7})$$

$$= \max_a \sum_f w_f \mu_{a,f}. \quad (\text{B8})$$

(B8) follows from $\mathbb{E}[x_{a,f}|b] = \mu_{a,f}$ and the linearity of expectation (c.f. Equation 7).

Marginal transition function. The marginal transition function is defined

$$T_{\text{meta}}(b' | b, c) = \mathbb{E}_{s \sim b} [T_{\text{meta}}(b' | b, c, s)]. \quad (\text{B9})$$

Unfortunately, it is not possible to simplify this expression in the general case. Turning to the multi-attribute case, recall that the transition function can be defined in generative form (rather than with an explicit transition probability function) as setting $\mu'_{a,f} = x_{a,f}$ and $\sigma_{a,f} = 0$. We want to create a similar generative model to produce b' given b . Because each

computation only updates the belief for one feature value, we can leave the others as is. Furthermore, $\sigma'_{a,f}$ does not depend on state, and so we can leave it as 0. This leaves $\mu'_{a,f}$. Here, we must account for our uncertainty in the true feature value, $x_{a,f}$. Because the computation sets $\mu_{a,f}$ to $x_{a,f}$, we can simply replace $x_{a,f}$ in the full transition function with a distribution capturing our belief about the value of $x_{a,f}$. By definition, $x_{a,f}$ is distributed $\text{Normal}(\mu_{a,f}, \sigma_{a,f})$. The complete marginal transition model is thus given by

$$\begin{aligned} \sigma'_{a,f} &= 0 \\ \mu'_{a,f} &\sim \text{Normal}(\mu_{a,f}, \sigma_{a,f}) \end{aligned} \quad (\text{B10})$$

with all other variables left unchanged.

One-step lookahead

The meta-greedy policy selects computations under the assumption that it will terminate computation on the next time step (if it does not terminate on this time step). Given this assumption, it selects the computation with maximal expected value. That is,

$$\pi_{\text{greedy}}(b) = \arg\max_c Q_{\text{greedy}}(b, c), \quad (\text{B11})$$

where Q_{greedy} denotes the one-step lookahead value. For the termination operation, there is no next step to look ahead to, so the policy uses the true expected value,

$$Q_{\text{greedy}}(b, \perp) = r_{\text{meta}}(b, \perp) \quad (\text{B12})$$

We have already derived an analytic expression for the expected termination reward in (B8). For all non-terminating computations, the the expected value marginalizes over possible outcomes of the computation (that is, the updated belief, b'):

$$Q_{\text{greedy}}(b, c) = \mathbb{E}_{b' \sim T_{\text{meta}}(b, c)} [r_{\text{meta}}(b', \perp)] - \text{cost}(c) \quad (\text{B13})$$

We now define an analytic expression for (B13). We begin by replacing r_{meta} with (B8), giving us

$$Q_{\text{greedy}}(b, c) = \mathbb{E}_{b' \sim T_{\text{meta}}(b, c)} \left[\max_a \sum_f w_f \mu'_{a,f} \right] - \text{cost}(c). \quad (\text{B14})$$

To ease notation, we introduce the shorthand $V_b(a) = \sum_f w_f \mu_{a,f}$ to denote the expected value of action a given belief b . The expectation then becomes

$$\mathbb{E}_{b' \sim T_{\text{meta}}(b, c)} \left[\max_a V_{b'}(a) \right] \quad (\text{B15})$$

Next, note that each computation only updates the expected value of a single action. We can thus split up the maximization into one part that depends on the updated belief and one that does not,

$$\mathbb{E}_{b' \sim T_{\text{meta}}(b, c)} \left[\max \left\{ V_{b'}(a_c), \max_{a \neq a_c} V_b(a) \right\} \right], \quad (\text{B16})$$

where a_c is the action inspected by computation c . Note that $V_{b'}(a) = V_b(a)$ for $a \neq a_c$ because computations only affect the expected value of one action. Thus, the internal max term is a constant with respect to the expectation. $V_{b'}(a_c)$, however, is a random variable. Specifically, it is Normally distributed with mean $V_b(a_c)$ and standard deviation $w_{a_c, f_c} \sigma_{a_c, f_c}$, with f_c denoting the feature inspected by computation c . This follows from

$$V_{b'}(a_c) = \sum_{f \neq f_c} w_f \mu_{a_c, f} + w_{f_c} \mu'_{a_c, f_c}. \quad (\text{B17})$$

as well as $E[aX+b] = b+a E[X]$ and $\text{Var}[aX+b] = a^2 \text{Var}[X]$. Intuitively, the expected value of the option should on average be the same after learning one of its features, and the size of the update (that is, the variance of the new expected value) depends on both the range of likely feature values, σ_{a_c, f_c} , and the amount the feature matters, w_{a_c, f_c} .

Thus, (B16) is the expected maximum of a Normally distributed variable, $V_{b'}(a_c)$, and a constant, $\max_{a \neq a_c} V_b(a)$. We can thus apply

$$E[\max\{X, z\}] = \Pr[X \leq z] \cdot z + (\Pr[X > z]) \cdot E[X | X > z]. \quad (\text{B18})$$

substituting $V_{b'}(a_c)$ for X and $\max_{a \neq a_c} V_b(a)$ for z . To write this expression in a compact and intuitive form, let $V_c = V_{b'}(a_c)$ be the value of the considered option after consideration (a random variable), let $v_{\text{other}} = \max_{a \neq a_c} V_b(a)$ be the value of the competing “other” option (a constant). We can then write

$$Q_{\text{greedy}}(b, c) = \Pr[V_c \leq v_{\text{other}}] \cdot v_{\text{other}} + \Pr[V_c > v_{\text{other}}] \cdot E[V_c | V_c > v_{\text{other}}] - \text{cost}(c). \quad (\text{B19})$$

This expression involves the Normal CDF and the expectation of a truncated Normal, both of which are provided by standard statistical libraries.

Appendix C

Results without exclusions

Here, we provide results for Experiments 1-3 without excluding participants who chose randomly (without clicking) on more than half of control trials. In the main text, we report the one case in which there is a difference in the significance of a predicted effect (problem complexity in Experiment 2). Here we report the full set of statistics.

Experiment 1

Participants chose the basket with the most prizes on 89.3% of trials when it was presented as the default op-

tion, compared with 51.2% on control trials. This difference was significant, as revealed in a logistic regression predicting default-chosen from nudge-present ($z = 43.61$, $p < .001$).

In the 18.0% of trials on which our participant did not immediately choose the default, they were still more likely to choose it eventually (63.0% vs. 57.6%; $z = 3.35$, $p < .001$).

As predicted, we found significant positive interactions with many-options ($z = 6.77$, $p < .001$) and many-features ($z = 1.65$, $p = .050$), and a significant negative interaction with idiosyncrasy $z = -4.80$, $p < .001$).

Participants achieved higher net earnings (payoff minus click cost) when a default option was presented (174.01 points vs. 161.50 points; linear regression: $t(12798) = 20.96$, $p < .001$), but there was a significant negative interaction between nudge-present and idiosyncrasy ($t(12796) = -3.06$, $p = .001$).

Experiment 2

Participants chose suggested options significantly more often than chance, as measured by a chi-square test of independence (38.0% vs. 16.7%, $\chi^2(1) = 2621$, $p < .001$).

As stated in the main text, participants were not significantly more likely to choose the suggested option in problems with five vs. two features, although the results trend in that direction ($z = 1.35$, $p = .089$).

In line with our predictions, participants were more likely to choose the suggested option when it was presented before an initial choice ($z = -10.75$, $p < .001$). However, we did not observe a significant interaction with problem complexity ($z = -1.55$, $p = .061$).

Experiment 3

Participants revealed an average of 2.13 values of the highlighted feature on nudge trials, compared with 1.13 values for control trials (two-sample t-test: $t(4009.8) = 16.42$, $p < .001$).

Participants chose baskets with an average of 5.87 prizes of the highlighted type, compared to a baseline of 5.52 on control trials ($t(4193.0) = 6.23$, $p < .001$). Similarly, participants chose the basket with the highest number of highlighted prizes significantly more often (53.3% vs. 44.2%; $\chi^2(1) = 35$, $p < .001$).

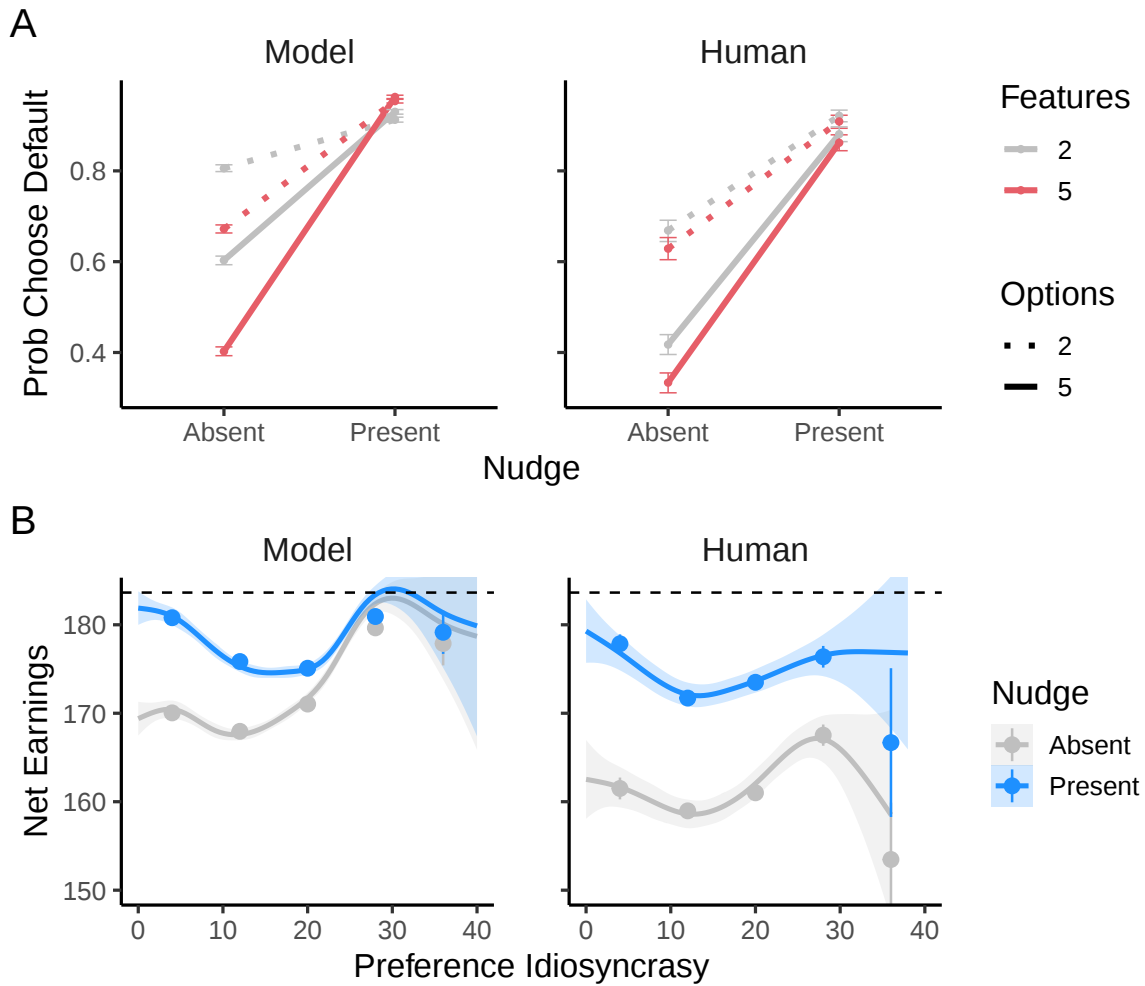
Figure C1*Experiment 1 results without exclusions**Note.* See Figure 7 for details.

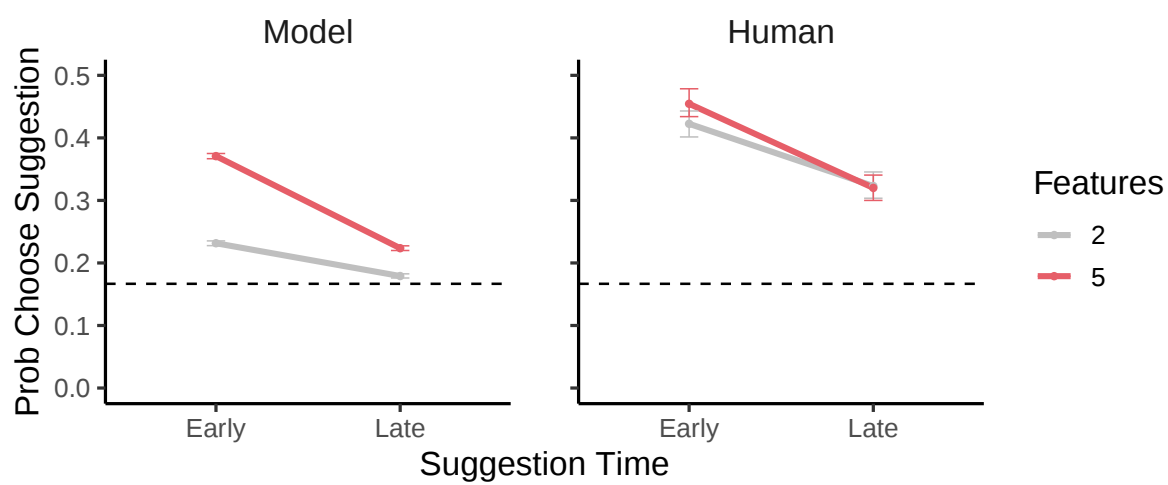
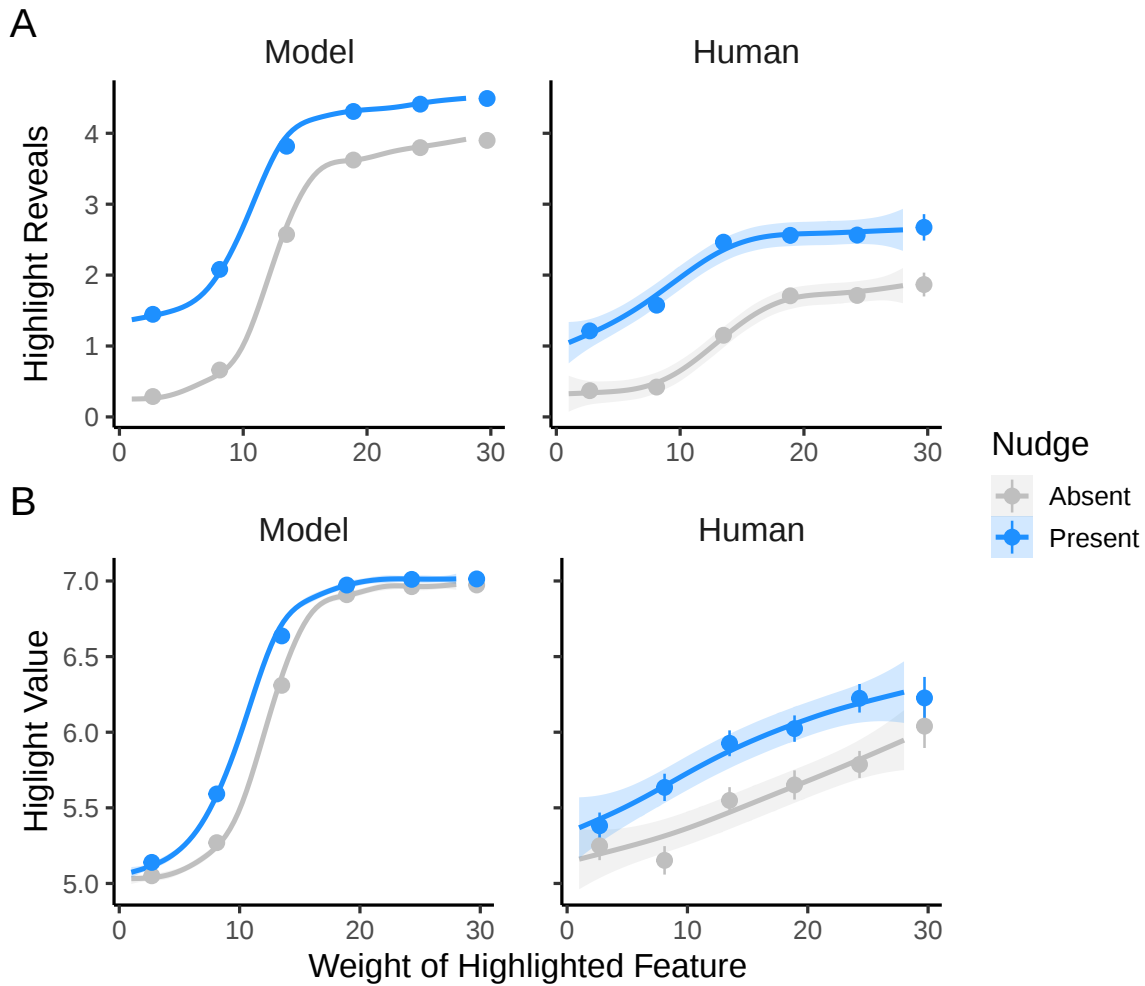
Figure C2*Experiment 2 results without exclusions**Note.* See Figure 10 for details.

Figure C3*Experiment 3 results without exclusions**Note.* See Figure 13 for details.