A USER'S GUIDE TO DEBIASING

Jack B. Soll, Katherine L. Milkman, and John W. Payne

INTRODUCTION

Imagine that you are wading through a large number of applications from prospective graduate students. Only several can be offered admission to the elite program at your university. Research shows that the type of decision you face is fraught with the potential for errant or "biased" judgments. For example, you may focus too much on the emotionally resonant stories that applicants share in their essays and too little on their qualifications. Alternatively, you may rely too heavily on stereotypes associated with certain college majors or on applicants' similarity to recently successful students in your program. Given that your objective is to choose the very best candidates—those most likely to stay in academia and do groundbreaking research—you would like to know how to avoid succumbing to these and other errors. Of course, some imperfect decisions are inevitable, but decades of research into the systematic biases known to plague decision makers have yielded an array of insights about strategies that can reduce the risk of bias. The purpose of this chapter is to provide a guide to these strategies. Our hope is that the ideas in this chapter can immediately be applied, so readers with some knowledge of judgment and decision research can go out straightaway and "do some debiasing." Naturally, there is still much research left to do, so we also hope that our discussion will prompt future work in this important area.

Before proceeding further, it is important to define what we mean by "debiasing". Kenneth Hammond (1996) introduced a fundamental distinction between coherence and correspondence criteria for evaluating the quality of judgments and decisions. Coherence criteria push judgments up against normative standards of internal consistency, such as Bayes' rule or transitivity of preferences. People are declared to be biased if they deviate in a systematic direction with respect to the normative standard. In contrast, correspondence criteria merely require the decision maker to be accurate but are unconcerned with questions of whether the decision maker is rational in the normative sense. The correspondence concept can also be expanded to evaluate a decision maker's welfare as defined by goal attainment, happiness, or satisfaction (cf. Dunwoody, 2009). We take an expanded correspondence view of debiasing in this chapter.

We will begin our discussion of debiasing judgment with a review of what is known about when people have the capacity for making good judgments and choices, a state we call *decision readiness*. Before thinking about improving judgments, we argue it is important to consider whether a decision should be made at all. Intense emotional states, fatigue, and poor decision-related skills (e.g., being innumerate) can all contribute to a lack of decision readiness. Next we will turn to a discussion of two general approaches available for debiasing decisions: (1) debiasing by modifying the decision maker (e.g., through education and the provision of tools) and (2) debiasing by modifying the environment (e.g., by creating optimal conditions to support wise judgment).

The first general approach to debiasing, which we dub "modify the decision maker," seeks to provide people with some combination of knowledge and tools to help them overcome their limitations and dispositions. This approach draws upon classic debiasing research on the benefits of education as well as thinking strategies, rules of thumb, and more formal decision

aids that people can be taught to use. For example, in the case of selecting graduate students, by developing a set of uniform inputs (e.g., grade point average, standardized test scores, quality of essay, etc.) for evaluating each candidate and assigning decision weights to those inputs (e.g., 10, 5, 10), you may be able to improve the accuracy and consistency of your evaluations.

The second approach, which we call "modify the environment," seeks to alter the setting where judgments are made in a way that either encourages better strategies or is a better match for the decision strategies that people naturally apply. This approach accepts that there is bias, but strives to create situations in which a bias is either irrelevant or may even be helpful. Prompted by Richard Thaler and Cass Sunstein's (2008) book *Nudge,* recent years have witnessed an explosion of research on choice architecture. Research in this area emphasizes improving decisions not by reforming people, but by reforming the contexts in which we learn, generate ideas, and make choices. For example, in structuring the environment where you evaluate prospective graduate students, a wise choice architect might remove students' photographs (which still appear at some institutions) before sending you applications to review. This would eliminate the opportunity to discriminate on the basis of appearance in forming impressions. Thaler and Sunstein make the important point that there is no such thing as a neutral architecture. Given this, why not set things up so that the best possible decisions are likely to ensue?

We begin this chapter with a brief discussion of the sources of bias in decision making. It helps to know how poor decisions arise in order to generate insights about how to improve them. This discussion is followed by a section on decision readiness, which both reviews the literature on when individuals are decision ready as well as suggesting what to do when they are not. The remainder of the chapter summarizes a variety debiasing techniques. These are organized into those that equip people with effective tools so that they are better adapted to their environment, and those that modify the environment so that it is better adapted to the way that people naturally make decisions.

SOURCES OF BIAS

Narrow Thinking

A problem for many decisions is that people form judgments based on the limited information that comes to mind or delivered by the environment, which is often incomplete or biased—a phenomenon that Kahneman (2011) calls WYSIATI (*What You See Is All There Is*). As an illustration, consider a study by Samuel Bond, Kurt Carlson, and Ralph Keeney (2008). The researchers asked MBA students to list all of the objectives relevant to choosing a summer internship after their first year in the program. Combining the objectives generated across participants led to the development of a master list including 29 objectives, such as "improves my attractiveness for full-time job offers," and "helps me develop my leadership skills." Students were next shown the master list and asked to indicate which objectives were important. They selected nearly fifteen objectives on average, but of these, only about seven were on their original, self-generated lists. In short, the students indicated that in choosing an internship, there were eight objectives they recognized as important but failed to think of on their own. Moreover, when asked to evaluate objectives' importance, the self-generated objectives were only rated slightly higher than others. This study powerfully demonstrates that people typically make decisions without developing complete criteria. In a follow-up study, Bond, Carlson, and

Keeney (2010) asked participants why they failed to uncover so many important objectives. Some mentioned *shallow thinking*—they simply devoted too little effort to the task. About half the participants chalked it up to *narrow thinking*—they focused their attention on one category of objectives, which crowded out their ability to identify other categories.

Some version of narrow thinking underlies many cognitive biases. In problem solving, people often become fixated on one type of solution, which impedes their ability to generate other, more creative ones. They may continue to generate new ideas, but these are often variants on a theme rather than truly novel (Smith, 2003). Narrow thinking also contributes to many classic biases in behavioral decision research including overconfidence, myopia, escalation of commitment, and anchoring, which are just several of many examples. Consider the anchoring bias for a moment. When estimating a figure such as how long it will take the Dow to reach 18,000, people are typically overly influenced by a starting value. This anchor can be a value posed in a question (e.g., "will it take more or less than 12 months for the Dow to reach 18,000?"), another person's opinion, or even an arbitrary number generated at random (Tversky & Kahneman, 1974). Anchors are influential because they guide processes of search and retrieval in the direction of anchor-consistent information, leaving obscured from view information that favors other possible answers (Chapman & Johnson, 2002; Strack & Mussweiler, 1997). Related effects have been observed when people make choices. In the process of choosing, people tend to interpret ambiguous information in a manner that favors the option towards which they are tentatively leaning (Carlson, Meloy, & Russo, 2006). Together, this research highlights that many successful debiasing strategies may succeed by broadening our thinking (Larrick, 2009).

System 1 versus System 2

The issue of narrow versus broader thinking is related to the processes by which we think. Many contemporary models of cognition conceive of the mind as composed of two overarching yet interconnected sets of processes. One set of processes is fast, effortless, and unavailable to conscious introspection (often referred to as *System 1*), whereas the other is slow, effortful, conscious, and can only process information sequentially (often referred to as *System 2*). Throughout this chapter, we adopt the common shorthand *System 1* and *System 2* to refer to these two sets of processes (Stanovich & West, 2000; Kahneman, 2011). Although System 1 evokes an image of mindlessness and being on "autopilot", it is critical to survival. System 1 excels at pattern recognition and works by association, meaning that it can grasp the essence of a situation and identify appropriate responses. The automatic output of System 1 has been theorized to arrive with an affective tag—a feeling about the goodness or badness of a stimulus (Ben-David, Graham, & Harvey, 1994; Slovic, Finucane, Peters, & MacGregor, 2007). System 1 is essential for recognizing safe foods, avoiding dangerous animals, and behaving appropriately in social situations.

Despite its importance, System 1 is susceptible to a number of systematic biases, including narrow thinking. Most biases pertaining to statistical reasoning such as base rate neglect, insensitivity to sample size, representativeness, the availability heuristic, and the conjunction fallacy are at least in part attributable to an over-reliance on the intuitive reactions of System 1 (Kahneman, 2011). Moreover, System 1 is largely responsible for present-bias—the tendency to overweight immediate gratification and underweight delayed benefits (Milkman, Rogers, & Bazerman, 2008).

A key function of System 2 is to monitor System 1, identify potentially incorrect responses, and take corrective action when needed. For example, imagine a relative approaches you with an investment opportunity in e-commerce. The idea strikes you as exceedingly creative, and evokes in your mind favorable comparisons with Amazon and eBay. Similarity assessment is the purview of System 1, and you will likely feel optimistic and confident about the project. However, this line of thought neglects important information, such as the fact that most new businesses fail. It is the job of System 2 to notice this omission, suspend judgment for long enough to think deliberatively, and to engage in appropriate statistical reasoning about your relative's realistic chances of success. In this case, an estimate of success that falls near the base rate for new e-commerce businesses is probably most appropriate (Kahneman & Lovallo, 1993).

Notice that poor judgment can arise at different points in a decision process. For example, System 1 might generate a biased response, System 2 might not notice a System 1 error, or System 2 might actually be the source of failure, applying the wrong principle. This framework suggests multiple avenues for debiasing. First, System 1 can be trained to generate better intuitions in the first place. Much of what is called expertise can be attributed to quickly assessing situations and generating correct responses. For example, experienced firefighters can often sense hazardous situations, without necessarily being able to articulate the cues that tipped them off (Klein, Calderwood, & Clinton-Cirocco, 1986). This level of intuitive performance can only be achieved by combining substantial experience with quick, unambiguous, and complete feedback—extended exposure to what Robin Hogarth (2001) calls a kind learning environment (see also Kahneman & Klein, 2009). Second, people can be taught cognitive strategies and correct principles (e.g., derived from economics or statistics), such that System 2 is more likely to generate a good response. Third, the monitoring of System 1 can be improved, either by sharpening the oversight function of System 2, or through environmental modifications that help provide the oversight.

DECISION READINESS

As described above, System 2 plays a critical role in decision making: it monitors judgments and, when necessary, corrects them. This is true even for experts, for whom strong System 1 associations based on contextual cues can occasionally lead to glossing over details that are atypical but important (Chi, 1986). We call a person "decision ready" when System 2 is capable of performing its functions, in terms monitoring, suspending decisions, and correcting judgments.

We highlight three determinants of decision readiness: fatigue and distraction effects that temporarily limit capacity, visceral influences related to one's physical and emotional state, and individual differences in ability and thinking style. The first two factors vary within a person—an individual may be more or less susceptible to the effects of emotion and fatigue at different points in time. The third factor varies across people—individuals who are better educated and more reflective are likely to outperform others when making certain types of judgments.

Fatigue and distraction. When concentrating hard, overriding natural impulses or making a series of difficult decisions, people become fatigued and depleted, which temporarily constrains the monitoring function of System 2. Tasks that require effort and attention such as exercising self-control, even if unrelated to the source of depletion, become more difficult (Baumeister et al., 2007; Milkman, 2012). System 2 can also be hampered by distraction or pre-occupation.

For example, when asked to hold a fact in memory (e.g., a seven-digit number), people are more likely to succumb to temptation (Shiv & Fedorikhin, 1999) and rely on activated stereotypes (Gilbert & Hixon, 1991). Time pressure is another temporary condition that can impede deliberative thought, in that it redirects attention toward finding quick solutions and can lead to greater reliance on System 1 (Payne, Bettman, & Johnson, 1993). Financial pressure (e.g., poverty) can also pull attention toward pressing needs, resulting in diminished cognitive capacity to reflect on far-sighted choices. This in turn may contribute to the challenges associated with climbing out of poverty (Shah, Mullainathan, & Shafir, 2012).

Visceral influences. Visceral reactions are essential to survival: recoiling in fear from a snake before System 2 has time to evaluate the situation in a reasoned manner is a wise move. However, there are many ways in which emotions and other visceral experiences can negatively impact decisions. People are prone to an empathy gap—they systematically overestimate the extent to which present desires will remain the same in a different physical or emotional state (Loewenstein, 1996; Loewenstein, O'Donoghue, & Rabin, 2003). For instance, hungry consumers select higher calorie snack options for next week, even when they know that the snacks will be consumed following a meal (Read & Van Leeuwen, 1998). Similarly, sexual arousal has a marked influence on attitudes toward risky or immoral behaviors (Ariely & Loewenstein, 2006). Importantly, even incidental emotions—those for which the source is unrelated to the task at hand—can temporarily distort beliefs and preferences (Lerner, Small, & Loewenstein, 2004; Gino & Schweitzer, in press).

Individual Differences. People differ in their decision readiness as a function of their intelligence, training, and thinking styles. Some biases, such as overconfidence and hindsight bias, correlate with cognitive ability, but many others do not, such as anchoring and attending to sunk costs (Stanovich & West, 2008). The tendency to engage System 2 appears to be more of a thinking disposition as opposed to an ability. More reflective individuals (i.e., those who more closely monitor their System 1 responses) are more likely to detect situations in which more careful reasoning is needed. They are also more likely to keep intuitive responses at bay while they take the time to do necessary additional thinking (Frederick, 2006; Toplak, West, & Stanovich, 2011). One pernicious problem is that even when a person is suspicious of System 1 output, she may not know the correct normative principles or algorithms to apply in order to improve on System 1. Stanovich and West (2008) call this a *mindware gap*. For example, many people lack the basic numeracy skills essential to making informed decisions about savings and debt (Lusardi & Tufano, 2009).

Decision readiness may be low due to a temporary state, such as hunger or distraction, or a more permanent condition, such as lack of training in normative rules (e.g., how to calculate conditional probabilities or compound interest). One way to improve decision making is to simply avoid making important decisions when depleted, fatigued, angry, aroused, hungry, or distracted. People can be advised to defer important decisions until they are in a sufficiently decision-ready state. However, this prescription begs the question: Can people recognize their own unreadiness, and do they know that it impedes cognitive ability? Although we are not familiar with any studies that address this question directly, research on self-awareness suggests that we should be pessimistic. People have a blind spot for their own biases, even though they often detect those same biases successfully in others (Pronin, Gilovich, & Ross, 2004). Simply

advising people to watch out when tired or moody is probably harmless, but the benefits may be small if recognizing our own decision readiness is challenging.

Modifications to decision environments may be a more promising approach to increasing decision readiness. Consider the case of Israeli judges making a series of parole decisions. A study by Danziger, Levav, and Avnaim-Pesso (2011) showed that Israeli judges were far more lenient immediately following meal breaks but became harsher as the time elapsed since their last break increased (denying parole requests with higher frequency). Parole applicants were clearly treated differently depending on when their cases came up. This inconsistency and arbitrariness could arguably be cured with a simple environmental modification—scheduling more breaks to nudge judges toward readiness. Some modifications reduce the penalties associated with unreadiness, as opposed to making sure that people are ready. For example, the Federal Trade Commission's Cooling-Off Rule gives consumers three days to cancel most purchases over $25 made in their homes under time pressure from door-to-door salesmen (Federal Trade Commission, 1996).

Many of the debiasing techniques described in the following sections of this chapter address unreadiness in one way or another. The techniques differ in which source of unreadiness they address, and whether they attempt to increase readiness or modify the environment so that readiness matters less.

MODIFY THE PERSON

Education

One way to improve decision making is to teach people appropriate rules and principles, to bridge the mindware gap. Students with coursework in economics or statistics are more likely to successfully apply fundamental principles from those disciplines to avoid biases. For example, such students are more likely to ignore sunk costs (Fennema & Perkins, 2008; Larrick, Mannes, & Soll, 2012) and to doubt the implications of small samples (Lehman & Nisbett, 1990). Relatedly, economics professors, as opposed to those in the humanities or biology, are more likely to apply economic principles in life, such as ignoring a sunk cost by leaving a disappointing movie early (Larrick, Nisbett, & Morgan, 1993). The findings are backed by laboratory studies showing that people can learn to apply statistical principals and transfer this learning to new domains, although transfer tends to fade over a span of weeks (Fong & Nisbett, 1991).

Education makes a difference, but how much? Do economics professors always ignore sunk costs, or just occasionally? Are statistics professors substantially less influenced by a first impression of a job candidate in an unstructured interview? If individuals are to debias themselves, they must not only possess the correct mindware, but must also identify the situations in which to apply it and be motivated to do so. Craig McKenzie and Michael Liersch (2011) provide an interesting demonstration of what it takes to debias the incorrect belief that investments grow linearly as opposed to exponentially. This bias causes people to underestimate the opportunity cost associated with postponing saving for retirement. Simply handing students calculators did not solve the problem; they plugged in the wrong math. Moreover, participants who knew about compound interest and understood how it works, as determined by a post-experiment survey, came no closer to determining the correct answers. Apparently, knowing the normative rule does not guarantee using it. What worked best were graphical interventions (a

modification of the environment) designed to vividly demonstrate the long-term benefits of saving now, and the penalties associated with delay.

Perhaps the most effective type of education is domain-specific training on a decision task that people will engage in repeatedly. For example, probability judgments are typically mis-calibrated—when judges are 90% sure that an event will occur, for instance, the actual rate of occurrence is typically far lower. Meteorologists, however, are surprisingly well-calibrated at predicting precipitation, and in general people can become well-calibrated with extensive training and feedback. Unfortunately, such training does not transfer easily. When well-calibrated experts switch to topics other than the ones on which they have been trained, they are as mis-calibrated as the rest of us (Keren, 1991).

None of this is to say that general education does not contribute to better decision making. There is a strong connection between poor mathematical ability (innumeracy) and susceptibility to certain decision errors, such as attribute framing (e.g., evaluating a test score of 74% correct as more favorable than 26% incorrect) and the ratio bias (Peters et al., 2006). More numerate individuals can also more accurately interpret and manipulate numerical information, such as that found in risk disclosures (Nelson, Teyna, Fagerlin, Lipkus, & Peters, 2008) and credit card statements (Soll, Keeney, & Larrick, 2011). Although modifications to the environment such as providing transparent disclosure can alleviate some of the costs of being innumerate, basic quantitative skills, such as those that a good education can impart, are critical as well.

Cognitive Strategies to Overcome Narrow Thinking

We all know at some level that it helps to look at problems from multiple perspectives. Yet people frequently do not do this when making decisions. It is perhaps no surprise, therefore, that one of the most successful debiasing techniques for tackling narrow thinking is to instruct people to look at a problem they face in another way. Although simple, this general approach needs to be adapted to the context. In this section, we give several examples of how the rule "look at it differently" can be applied successfully to different decision-related tasks.

Generating Alternatives. Although much decision research has focused on whether people make optimal choices given a fixed set of options, having a good set of alternatives from which to choose is at least as important as choosing wisely. Past research suggests that when generating alternatives, clear decision objectives should be relied upon (Keeney, 1992; Kelley, 2001). Drawing on both field experience and experiments, Keeney (2012) concludes that people generate more alternatives when their decision objectives are considered one at a time rather than all-at-once. By focusing on objectives sequentially, a decision maker is able to adopt a new perspective with each iteration of the alternative generation process, which is likely to lead to the generation of a diverse set of options covering multiple categories of solutions. This process can potentially backfire if each generated alternative focuses too narrowly on a single objective to the detriment of others. Another round of alternative generation may be needed to create hybrid alternatives that better satisfy the complete list of objectives. Nevertheless, the strategy of taking on objectives one at a time appears to help decision makers identify categories of alternatives that otherwise would run a higher risk of being ignored.

Dealing with Optimism. On any new undertaking, people tend to overestimate their chances of success (Moore & Healy, 2008). This is the result of our tendency to focus too narrowly on evidence that supports our initial hypotheses and to underweight contradictory evidence. One fix is to simply take the time to consider reasons why we might be wrong or fail (Koriat, Lichtenstein, & Fischoff, 1980). Russo and Schoemaker (in press) discuss a modified version of this approach, *prospective hindsight*, which might be even better. To apply this strategy, imagine time-traveling into the future and learning that your undertaking has failed. For example, a prospective home buyer in the year 2013 might ask "Here in 2033, why is my house worth less than what I paid for it twenty years ago?"). When contemplating a past failure, even if only imaginary, people tend to identify potential causal paths that do not come to mind in foresight (Mitchell, Russo, & Pennington, 1989). Although more research is needed, in principle, prospective hindsight should help dampen excessive optimism as well as helping people better plan for different possible contingencies.

Improving Judgmental Accuracy. Stefan Herzog and Ralph Hertwig (2009) proposed a clever technique called *dialectical bootstrapping* to increase the amount of information brought to bear on quantitative judgments. After an initial best guess of a quantity (e.g., the Dow's 2016 year-end close), assume that you are wrong, guess again, and then average the two guesses (do not choose your favorite!). The two guesses may be very different from each other, because by design they are supposed to represent different perspectives or subsets of your knowledge. Two such guesses will therefore often bracket the truth, leading errors to cancel out in many cases when the guesses are averaged. The average guess generated in this manner tends to be more accurate than always sticking with your first or second guess, and also more accurate than choosing the guess believed to be most accurate. Averaging guesses from two or more people is even better. Even so, using techniques like dialectical bootstrapping to harness "the crowd within" achieves about half of the accuracy gain that could be achieved by averaging guesses from two different people (Larrick, Mannes, & Soll, 2012; see also Vul & Pashler, 2008).

Assessing Uncertainty. One of the most robust forms of overconfidence arises when people provide assessments of their confidence intervals (e.g., "I am 80 percent sure that the house will sell for between 250 and 275 thousand dollars"). For example, over a nine-year time horizon, the researchers Ben-David, Graham, and Harvey (2010) asked corporate chief financial officers to forecast yearly returns for the S&P 500. Although the CFOs presumably had vast knowledge of the U.S. economy, their 80% intervals captured the true answers only 33% of the time, implying that they were far too often surprised by the outcomes.

Although it is difficult to completely erase this type of overconfidence, two methods of debiasing have proven helpful. The first method is to split the question into multiple parts that force judges to focus separately on low, medium, and high answers. For example, asking for the 10^{th}, 50^{th}, and 90^{th} percentiles of a subjective interval distribution improves the hit rate by about 20 percentage points compared to asking a single question requesting one range. This method produces intervals that are both wider and better centered on the truth (Soll & Klayman, 2004; see Haran, Moore, & Morewedge, 2010, for a related method). The second method, time unpacking (Jain et al., in press), applies to forecasting time series. Rather than forecasting the price of the gold three months in advance, for example, this method has forecasters assess intervals for one and two months in the future before producing a three-month interval forecast. For three-month horizons, time unpacking increases hit rates by about ten percentage points.

Time unpacking gives forecasters the sense that they are forecasting further into the future, leading them to feel more uncertain about their estimates and thus to provide wider (and therefore better calibrated) confidence intervals.

Using Models to Decide

Quantitative Models. One of the most straightforward and well-validated means of debiasing judgment is to take it out of the equation altogether, or rather, to replace it with an equation. As recounted in the bestselling book *Moneyball* (Lewis, 2003), equations have revolutionized the market for professional baseball players. This market was polluted for years by expert judgments that overweighted available information (e.g., recent performance and perceived similarity to other players), and failed to properly account for many of the most important predictors of a player's value (e.g., on-base percentage and ability to avoid strikeouts). Teams that began relying on equations built on valid predictors (rather than expert judgments) acquired a major performance advantage until their competitors began to also develop and apply equations.

Past research has shown that linear models outperform expert judgment across a wide range of settings (Dawes, Faust, & Meehl, 1989). The most sophisticated form of linear model requires historical data on relevant decision inputs (e.g., when admitting graduate students: GPA, GRE scores, strength of undergraduate university, strength of recommendation letters, etc.) as well as historical data on decision quality (e.g., student performance). Such historical data makes it possible to fit a linear model to characterize the relationship between various inputs and the output of interest. The resulting model suggests appropriate weights to place on various decision inputs when summing them in order to best forecast outcomes. Amazingly, even the simplest linear models that equally weight all decision inputs (thus requiring no historical data for calibration) outperform expert judgments (Dawes & Corrigan, 1974; Einhorn & Hogarth, 1975). Nearly as impressive is the performance of another type of linear model fitted to predict expert judgments (e.g., which baseball players are rated highly by scouts) rather than actual historical outcomes (e.g., player performance). Such bootstrapped models have been shown to outperform the very expert judgments they model, presumably by reducing the noise inherent in experts' decision rules (Dawes & Corrigan, 1974; Camerer 1981; Payne et al., 1993). The value of linear models is that they systematize the reliance on relevant decision criteria and eliminate the opportunity for human biases (e.g., the tendency to overweight recent information, the conjunction fallacy, gender bias) to creep into a decision and reduce its quality. When human judgment is critical to predicting outcomes, however, importantly, it can be entered as an input into a model.

Linear models are just one of many types of quantitative models that can help systematize judgments in order to reduce opportunities for error. See Chapter 40 in this volume for a discussion of other useful quantitative modeling tools.

Checklists. Closely related to quantitative models are checklists, which provide a particularly potent tool for streamlining processes and thus reducing errors (Gawande, 2010). A checklist provides "a list of action items or criteria arranged in a systematic manner, allowing the user to record the presence/absence of the individual item listed to ensure that all are considered or completed" (Hales & Pronovost, 2006). Checklists reduce errors due to memory failure and are most widely used in aviation (USAF Series, 1999) where the transition from paper-based to

electronic checklists to regulate pre-during-and-post-flight routines decreased errors by 46% (Boorman, 2001). In medicine, adopting checklists to guide doctors' and nurses' interactions with patients has been shown to reduce errors in fields ranging from trauma to anesthesiology (Runciman et al., 2005; Harrahill & Bartkus, 1990; Hart & Owen, 2005; Myburgh, Chapman, Szekely, & Osborne, 2005) as well as significantly improving the care of regular hospital inpatients (Wolff, Taylor & McCabe, 2004). Checklists are particularly valuable in settings where best practices are likely to be overlooked due to extreme complexity or under conditions of high stress or fatigue (Hales & Pronovost, 2006).

Importantly, models have some significant limitations. First, they are only as valuable as the attributes they include. If in building a quantitative model or checklist, important inputs are overlooked (e.g., due to narrow thinking), the solution produced may be biased. Second, under special or changing circumstances, models based on historical data will not apply and may even lead you wildly astray. At moments when such circumstances arise, human judgment is necessary and a deviation from strictly relying on solutions generated by models is required, but determining when and where models won't apply is one of the greatest challenges associated with relying on these valuable decision tools.

MODIFY THE ENVIRONMENT

Another, different approach to debiasing is to change the environment in which a decision will be made in ways that are likely to reduce the incidence of error. The first means of modifying the environment that we will discuss here includes approaches that rely on *economic levers* to change behavior for the better, such as increasing the incentives associated with wise choices or providing new information to clarify which choices are wisest. A second means of shifting the environment for the purpose of debiasing uses well-understood psychological principles as a tool to improve potentially biased decisions. Specifically, anticipating a common potential error (e.g., under-saving for retirement), someone designing a decision-making environment - a "choice architect" - can structure that environment (e.g., the procedure for enrolling in a 401k) to "*nudge*" choices in wise directions (Thaler & Sunstein, 2008). Note that sometimes people can modify the environment for their future selves—they can be their own choice architects.

In the remainder of this section, we will elaborate on a number of debiasing strategies that fall into the two categories of "make changes to the environment" described here.

ECONOMIC LEVERS

We begin by summarizing the benefits and limitations of relying on economic levers to change the environment in order to promote wise choices and thus reduce bias. Below, we describe some of the most widely-used economic approaches to environmental modification.

Incentivize Wise Choices. Incentives have the power to improve many decisions by increasing the costs of a bad choice and thus increasing decision makers' motivation to choose wisely. For example, when it comes to self-control problems, incentives can help people resist temptation: monetary incentives can successfully encourage weight loss (Volpp et al., 2008), drug adherence (Volpp et al., 2008), smoking cessation (Volpp et al., 2006), and exercise

(Charness & Gneezy, 2009; Acland & Levy, 2011). In settings where habit formation is key to long-term behavior change, providing incentives for repeated engagement in a behavior can be particularly beneficial by creating habits that last after incentives are removed (Charness & Gneezy, 2009; Acland & Levy, 2011). Unfortunately, when it comes to addressing cognitive biases, incentives have a more checkered history of success (Fischhoff, 1982). While monetary incentives typically improve performance on judgment tasks that can be tackled more successfully by increasing effort (e.g., memory/recall tasks), incentives have not been able to eliminate most rationality violations (Camerer and Hogarth, 1999). When effort and attention aren't sufficient to reduce errors in judgment, as is the case with many biases (e.g., hindsight bias, overconfidence), incentives do not offer a cure (Fischoff, 1982; Camerer & Hogarth, 1999). Importantly, monetary incentives can even *backfire* in some cases by increasing anxiety (e.g., choking in sports and test-taking) (Camerer, 1998; Baumeister, 1984) and by heightening individuals' tendencies to try to improve upon reliable formulas using idiosyncratic knowledge (Arkes, Dawes, & Christensen, 1986).

Introduce Accountability. Just as incentives can increase decision makers' motivations to choose wisely, making people accountable for their actions can have a similar motivational effect (Lerner & Tetlock, 1999). For example, one highly effective get-out-the-vote strategy is to inform an individual that a letter will be sent to all of her neighbors following an election listing each neighbor's (public) voting record (Gerber, Green & Larimer, 2008). Here, accountability helps people overcome their tendency to ignore civic duties in favor of more immediately gratifying pursuits, and it gets them to the polls. Like incentives, accountability primarily leads to increased performance on tasks where higher effort levels produce better outcomes (Simonson & Nye, 1992; Huber & Seiser, 2001; Lerner & Tetlock, 1999). Again, however, when effort is not the solution to debiasing, accountability does not improve outcomes. Further, because outward accountability heightens the desire to make decisions that appear internally-consistent and are easily to justify, accountability can increase self-presentation biases (Simonson, 1989).

Provide New Information. In some situations, providing new information or ratings to individuals can improve decisions. For instance, in the domain of self-control, a field study at Starbucks restaurants found that posting calorie information for foods and beverages tends to reduce the average calories purchased per transaction (Bollinger, Leslie & Sorenson, 2010). In theory, new information affects choices because it changes beliefs about the expected utility to be obtained from different options faced and thus can alter the optimal decision (Healy & Palepu, 2001). However, despite a policy maker's best intentions, providing information can sometimes make decision makers worse off. For one, information can be disheartening without improving decisions (e.g., consumers might enjoy meals less when calorie information is posted but may not gain health benefits from their new knowledge). Second, providing too much information can lead to decision paralysis due to information overload (Iyengar & Lepper, 2000). This highlights the need for presenting new information in a helpful manner, an important means of modifying the environment that we address in a later section of this chapter about kindly shaping information. Finally, for the special case of conflicts of interest (e.g., a financial advisor recommending mutual funds), advisors tend to act more in line with their own self-interest when they are required to disclose conflicts of interest because they feel licensed to behave selfishly, having fairly warned their advisee (Cain, Loewenstein & Moore, 2005). In general, providing

new information can be a helpful means of preventing errant judgments resulting from inadequate knowledge, although clearly, there are some caveats.

Convey Social Norms. Individuals have a tendency to herd, or to imitate the typical observed or described behaviors of others (Cialdini, Kallgren & Reno, 1991), in large part because the behavior of the herd typically conveys information about wise courses of action. This tendency can be used strategically in the provision of new information to improve decisions. For example, providing social norms information about the energy usage of one's neighbors on an electricity bill can sustainably reduce energy consumption by 2% (Alcott, 2011). Indicating that most hotel visitors re-use their towels boosts re-use rates (Goldstein, Cialdini, & Griskevicius, 2008), and similarly, emphasizing high anticipated turnout in elections increases voters' likelihood of going to the polls (Gerber & Rogers, 2009). Importantly, because people follow the crowd, providing social norms information can backfire when the message conveyed is that few people are engaging in a given beneficial behavior (Shultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007). Social norms messaging can also backfire when mustered in a challenging arena, as it can be demotivating to learn that the majority of others are far ahead when it will be hard to catch up (Beshears, Choi, Laibson, Madrian, & Milkman, 2012).

We have highlighted that standard economic approaches to debiasing can be effective in some cases, but they also have many limitations. Recent research suggests additional, valuable ways of modifying the environment to reduce errors that can be even more effective than standard economic approaches in some cases.

NUDGES

For a modification of the enviornment to qualify as a "nudge", the design change cannot restrict choice or alter prices but must instead make use of psychological principles to influence behavior for good (Thaler & Sunstein, 2008; Ly, Mazar, Zhao & Soman, 2013). An example of a nudge can help illustrate this concept: making enrollment in a company's retirement savings program the default (or "opt-out") is a nudge that can increase enrollment by as much as 37% (Madrian & Shea, 2001). Default rules don't restrict choice and thus won't influence those with strong preferences, but we know that people tend to exhibit inertia in the face of a default (Johnson, Bellman & Lohse, 2002). Thus, we can leverage this psychological principle to help people overcome a common decision-making pitfall: the tendency to under-weight future utility and exhibit present-bias. Below, we describe some of the most widely-used and widely-tested nudges that have been shown to debias judgment successfully across a number of settings.

Defaults

As described above, one of the most potent known nudges involves paternalistically setting the default option, or the choice that will be selected unless a decision maker actively opts out. Defaults are effective nudges because they leverage decision makers' inertia, suggest an implicitly recommended course of action, and set a status quo that loss aversion makes it painful to sacrifice (Johnson et al., 2002; McKenzie, Liersch, & Finkelstein, 2006). Past research has shown that in countries where organ donation is the default and citizens must opt-out of donating, donation rates are approximately 90 percentage points higher than in countries (like the

United States) where citizens must opt-in to become donors (Johnson & Goldstein, 2003). Defaults have also been employed successfully to increase flu vaccination rates (by defaulting employees into flu shot appointments) (Chapman, Li, Colby, & Yoon, 2010) and retirement savings (Madrian & Shea, 2001), among other things (see Goldstein et al., 2008). Unlike many nudges, defaults do not tackle a particular class of bias but instead can help individuals make better choices whenever they face the task of deciding between alternatives.

Nudges that Kindly Shape Information

People are more likely to reach accurate conclusions when they have the right information packaged in an intuitively comprehensible format. In principle, a sophisticated consumer could repackage information on her own. However, people often neglect to do this for a variety of reasons (e.g., it requires too much effort, they lack the required skills, or their System 2 fails to detect the necessity). Below we describe several nudging strategies that can be used to shape information so it will be particularly impactful for the purposes of debiasing.

Metric transformations. Consumers encounter many metrics of efficiency designed to assist with their purchasing decisions: MPG (miles per gallon) for vehicles, SEER ratings for air conditioners, and megabytes per second for data transfer. These metrics share a common property—the relationship to the variable relevant to the consumer's objective is nonlinear (minimize fuel consumption for the first two, and minimize time for the third). For example, imagine that you upgrade your old sedan that gets 25 MPG for a hybrid that gets 50. This seems like a very big improvement, especially compared to that achieved by your neighbor who traded in an old gas-guzzling pickup truck and increased his MPG from 10 to 15. Mile for mile, however, your neighbor will save more gas. Most people fall prey to this "MPG illusion" by failing to recognize that efficiency gains are greater on the low end of the scale (your neighbor will save 3.3 gallons for every hundred miles he drives, compared to your 2 gallons saved). Providing consumers with the inverse of MPG, or GPM (gallons per hundred miles), leads to more accurate perceptions of fuel efficiency (Larrick & Soll, 2008). For this reason, the EPA has added this new metric (GPM) to vehicle labels beginning with model year 2013. This example illustrates that providing information in the appropriate units given decision makers' objectives can improve decisions.

In some cases even a linear transformation can turn an unintuitive metric into a more relevant quantity by drawing attention to an attribute consumers care about (Camilleri & Larrick, 2013). For example, the new EPA vehicle label described above states how much money a consumer can expect to save in fuel costs over 5 years if she buys a given car rather than an average new vehicle. This metric just multiplies GPM by the expected price of gas and miles driven over 5 years, so from a pure information perspective it provides no added value. However, most consumers are unlikely to do this calculation on their own, even if they care about long-term operating costs. Importantly, in providing such transformations, the scaling factor selected is an important consideration, as it can affect the weight placed on an attribute (Burson, Larrick, & Lynch, 2009). The EPA chose to state costs over 5 years, but they could have chosen a different scale (e.g., 10 years, 1 year, or 3 months). Arguably, the 5-year time frame is appropriate because it matches the typical vehicle ownership period, places gas consumption in the context of other large purchases, and nudges people toward more environmentally friendly vehicles.

Kind Representations. For about twenty years the USDA used the Food Pyramid diagram as a visual guide indicating how much a typical American should eat from different food groups. Unfortunately, the food pyramid was too abstract to be a useful. According to Chip and Dan Heath (2011), "The Food Pyramid might as well be a Food Rhombus or a Food Rooster." The authors describe research showing that people are best able to change their behavior when they have a concrete model to follow, and the pyramid in no way provided such a model. The USDA's new MyPlate diagram, which has replaced the pyramid, appears to solve the problem by providing a more intuitive model. It is a picture of a plate showing how to divide meals across the food groups. Half the plate is filled with fruits and vegetables.

Intuitive representations are not only helpful when it comes to conveying guidelines. They can also help people reason more accurately about probabilities, such using base rates in accordance with Bayes' rule. Providing relative frequency information can improve judgment (e.g., 10 out of every 100 instead of 10%), but in order to be effective, such information needs to be conveyed in a manner that helps people understand the structure of the problem they face (Barbey & Sloman, 2007). One promising way to improve decision makers' comprehension of probabilities is with pictorial displays (Galesic, Garcia-Retamero, & Gigerenzer, 2009). For example, Fagerlin, Wange, and Ubel (2005) asked participants to choose between two procedures for heart disease—either bypass surgery with a 75 percent chance of success, or a less arduous procedure, angioplasty, with a 50 percent chance of success. Participants relied much less on irrelevant anecdotal information in making decisions when procedures' stated success probabilities were accompanied by 10x10 grids that visually displayed their success rates.

Smart Disclosure. Many major consumer decisions have become increasingly complex. When it comes to selecting a mortgage, a drug prescription plan, or a cell phone plan, consumers face a dizzying array of options that vary on many attributes. Merely improving disclosure may not be enough to guide people toward good choices and fend off deception by unscrupulous vendors (Thaler, 2012). According to a Presidential task force on the subject, "Smart disclosure requires service providers to make data about the full range of their service offerings available in machine-readable formats such that consumers can then use these data to make informed choices about the goods and services they use." The hope is that a market would evolve for third-party web sites to combine this information with consumers' stated preferences and personal data (e.g., cell phone usage, drug prescriptions, etc.) to provide tailored menus that dramatically simplify choice.

Nudges to Induce Reflection

A number of important decision biases emanate from an under-reliance on System 2 thinking (or over-reliance on System 1 thinking) and can thus be reduced by nudging deeper reflection. Such nudges require people to devote more time and attention to a decision, often by more explicitly elaborating their objectives and plans. Below, we discuss interventions that prompt additional thought.

Planning Prompts. Prompting the formation and articulation of concrete plans to complete a desired action is one form of induced reflection that can help decision makers avoid failures to follow-through on their intentions due to both procrastination and forgetfulness (both errors that

have be argued to emanate from failures of System 2). When an individual is prompted to decide on the when, where and how of an intended action, the plans formed become a commitment that is both (1) psychologically difficult to break and (2) memorable because of the link between the action and a concrete future moment (Gollwitzer & Sheeran, 2006). Planning prompts have been used successfully to increase many important outcomes including voter turnout (Nickerson & Rogers, 2010), flu shot take-up (Milkman, Beshears, Choi, Laibson, & Madrian, 2011) and colonoscopy rates (Milkman, Beschears, Choi, Laibson, & Madrian, 2013). They are particularly efficacious when plan-making is unlikely to come about organically (Nickerson & Rogers, 2010) and when forgetfulness poses a large obstacle to follow-through (Milkman et al., 2011, 2013).

Forced Breaks. Another way of reducing judgment errors resulting from an under-reliance on System 2 reasoning is to build forced breaks into choice environments in order to encourage added reflection. Specifically, researchers have shown that partitioning the consumption of food (e.g., by presenting the same quantity in a set of separate containers rather than one container) reduces mindless eating, and partitioning gambling tokens reduces gambling (Cheema & Soman, 2008; Delaney, 2007). By partitioning resources, natural breaks are imposed on choice processes at the moment when a partition is encountered, and these breaks lead to slower decisions and deeper processing. Such breaks are most valuable when imposed on decisions where an individual intends to regulate consumption but sometimes fails due to mindlessness (Cheema & Soman, 2008).

Active Choice. Requiring decision makers to actively choose between multiple options rather than simply avoiding a choice and accepting a default is another nudge towards induced reflection capable of improving certain judgments. Recent research has shown that compulsory choice helps decision makers avoid mindlessly accepting defaults that may not be ideal for them. For example, requiring prescription drug users to make an active choice between receiving medications at their local pharmacy versus by home delivery (at a discount) increases home delivery rates by 35 percentage points (Beshears, Choi, Laibson & Madrian, 2012). In the retirement savings domain, requiring new employees to make a compulsory choice about 401k enrollment increases enrollment rates by 28 percentage points over an opt-in choice scheme (Carroll, Choi, Laibson, Madrian & Metrick, 2009). These findings highlight that choice architects can use active choice requirements as a tool to prevent mindless acceptance of defaults from leading to biased judgments. However, because of the added cognitive tax imposed on decision makers when an active choice is required, this debiasing technique may be best saved for decisions of some consequence.

Checklists. It is worthy of note that in some cases where shallow thinking leads to errant decisions, as in the examples cited above, merely prompting System 2 thinking using the techniques described above is enough. These are situations in which people would know exactly what to think about, if only they stopped to think. In other cases, prompting deliberation is not quite enough, because the person would not know what to think about. For example, an airplane pilot needs to evaluate many details before deciding whether or not it is safe to take off. Certainly it is important that the pilot is not fatigued, to avoid the cognitive equivalent of autopilot. But even when engaging in careful thought, a pilot is bound to overlook essential details when operating a complex machine with no oversight. In such cases induced reflection

must not only guide people to think, but also direct the content of that thought. Earlier in this chapter, we discussed the value of checklists as a potent tool for streamlining processes and thus reducing errors (Gawande, 2010). Checklists can be viewed as a nudge towards greater reflection designed to guide thinking rather than merely to prompt it.

Nudges to Induce Future-Focused Thinking

Present-bias, or the tendency to overweight immediate gratification while underweighting the long-term implications of a choice, is arguably responsible for many errors in judgment. Specifically, present-biased thinking has been blamed for societal problems ranging from obesity to under-saving for retirement (O'Donoghue & Rabin, 1999; Milkman, Rogers & Bazerman, 2008). Below, we describe a series of nudges designed to promote future-focused thinking in order to reduce the pernicious effects of near-sightedness.

Choose in Advance. One means of reducing biases related to near-sightedness is prompting individuals to make decisions well in advance of the moment when those decisions will take effect (Rogers & Bazerman, 2008). This strategy is impactful because of people's tendency to make less impulsive, more reasoned decisions (relying more heavily on System 2) when contemplating the future than the present (Milkman et al., 2008). Choosing in advance has been shown to increase support for wise policies requiring sacrifices (Rogers & Bazerman, 2008), to increase charitable giving (Breman, 2011), to contribute to increases in retirement savings (Thaler & Benartzi, 2004),[1] and to increase the healthfulness of consumers' food choices (Read & van Lewun, 1998; Milkman, Rogers & Bazerman, 2010). Another result of choosing in advance is that choices are made in a higher construal level mindset, which means they focus more on abstract (e.g., why?) rather than concrete objectives (e.g., how?) (Trope & Liberman, 2003, 2010). This has other by-products, however—for example, higher construal leads to greater stereotyping (McCrea, Wieber & Myers, 2012). Therefore, an important caveat to choosing in advance is that it may lead to greater discrimination against women and minorities, as demonstrated in a field study of decisions on whether to grant prospective graduate students requests for meetings (Milkman, Akinola & Chugh, 2012).

Pre-commitment. Because people tend to make more patient and reasoned decisions about the future, as described above, providing opportunities for individuals to *both* choose in advance *and* make a binding decision (or at least decisions where penalties will accompany reversals) can improve many choices (Ariely & Wertenbroch, 2002). A classic example of pre-commitment comes from the story of Odysseus who tied himself to the mast of his ship (thus relinquishing control of his vessel) before encountering the enchanting music of the Sirens – sea nymphs famous for luring sailors towards deadly rocks. Research has shown many benefits from pre-commitment. For example, substantial savings increases result from providing individuals with

[1] In a seminal study showing the power of choice architecture for improving welfare, Thaler and Benartzi leveraged not only the power of choosing in advance but also loss aversion in order to dramatically increase employees' retirement savings. Specifically, they created a savings plan called "Save More Tomorrow™" that would (a) only require contributions in the future and (b) would never reduce an employee's salary because all contributions would be taken as a portion of an employee's future pay raises. Seventy eight percent of those offered the plan joined and their savings rates increased by 3.5-13.6% over the course of 40 months following enrollment (Thaler & Benartzi, 2004).

access to bank accounts with commitment features such as a user-defined savings goal (or date) such that money cannot be withdrawn before the pre-set goal (or date) is reached. Although only 28% of those offered such commitment accounts selected them when equivalent interest was available on an unconstrained account, average savings balances increased by 81 percentage points for those customers of a Philippine bank with access to commitment accounts (Ashraf, Karlan & Yin, 2006). Recent research has also shown that pre-commitment can help people quit smoking (Giné, Karlan, & Zinman, 2010), exercise (Royer, Stehr, & Sydnor, 2012), achieve workplace goals (Kaur, Kremer, & Mullainathan, 2010) and resist repeated temptation in the laboratory (Houser, Schunk, Winter, & Xiao, 2010). Pre-commitment is particularly valuable in settings where self-control problems pit our long-term interests against our short-term desires. When it comes to food, for example, pre-committing to smaller plates and glasses can reduce consumption (Wansink & Cheney, 2005).

Temptation Bundling. A new twist on pre-commitment called "temptation bundling" actually solves two self-control problems at once. Temptation bundling devices allow people to pre-commit to coupling instantly gratifying activities (e.g., watching lowbrow television, receiving a pedicure, eating an indulgent meal) with engagement in a behavior that provides long-term benefits but requires the exertion of willpower (e.g., exercising, reviewing a paper, spending time with a difficult relative). Such pre-commitment devices can increase engagement in beneficial behaviors like exercise while reducing engagement in guilt-inducing, indulgent behaviors (Milkman, Minson & Volpp, 2012).

Discussion

Nudges have proven to be particularly valuable in settings where bias can be anticipated and where a benevolent policy maker or manager has the opportunity to redesign the decision environment. In particular, they are a constructive way to combat forgetfulness, laziness, and the tendency to make near-sighted, present biased choices. However, importantly, in many cases, there is no benevolent policy maker with information about what an individual's "optimal" choice ought to be, rendering nudges an irrelevant tool.

GENERAL DISCUSSION AND CONCLUSIONS

We began this review of the debiasing literature by highlighting two prominent sources of bias – narrow thinking and under-utilization of System 2 thinking. We then turned to a discussion of "decision readiness", highlighting situations when decision makers are unprepared to make unbiased decisions, often due to constraints on System 2, but sometimes due to a mindware gap that can be traced to educational or intellectual deficits. Finally, we described two categories of paths to improving decisions: (1) modifying the person through either education or the provision of strategies and tools, and (2) modifying the environment a decision maker faces to facilitate wiser choices.

Our discussion has focused on strategies that can be used to debias individual judgment, but has not described a large and important literature suggesting strategies for improving the decisions made by groups. Although groups can potentially benefit from multiple minds tackling a problem from different perspectives, the actual performance of a group depends on group composition, group process, and the representation and distribution of knowledge across group

members. Depending on the interplay between these factors, groups can often improve upon individual judgment, but they can also amplify or introduce new biases (Kerr & Tindale, 2004).

We have also focused primarily on debiasing techniques that do not require providing individuals with extensive input or assistance from others. Recently, a burgeoning literature on advice-taking has revealed a central finding: people under-utilize advice (Soll & Mannes, 2011; Yaniv & Milyavski, 2007). Advice from others can help decision makers overcome narrow thinking, because a different person will likely bring a new perspective to a problem. However, this presumes that the advisor knows something different than the advisee, and is not anchored by the advice-seeker on an answer prior to giving advice (Larrick, Mannes, & Soll, 2012).

Increasingly governments around the world are looking to improve the decisions made by their citizens. On the plus side, many of the interventions discussed in this chapter offer the possibility of great benefits at relatively little cost, such as the work by Benarzti and Thaler on the Save-More-Tomorrow program. On the other hand, in the case of nudges they are also debating the ethics of interventions to change behaviors that, in many cases, operate below awareness. Related to this debate, an important issue for any debiasing tactic is whether it succeeds in making people better off. For example, nudges to increase savings may benefit most people, but those who should be spending more might be nudged in the wrong direction. Similarly, revised disclosure statements designed to improve comprehension might end up creating new sources of confusion as they correct old ones. Given the rapid development of new nudges in the literature, a consensus is needed regarding how to evaluate and measure success. For some initial progress in this area and an interesting discussion, see Ubel (2012).

On judgments ranging from allocating savings to retirement to hiring employees to assessing risks, bias is a common but not insurmountable human problem. As described at the outset of this chapter, we hope our review of the debiasing literature will better-equip readers with a set of strategies for improving decisions as well as stimulating future research on the important topic of debiasing.

REFERENCES

Acland, D., & Levy, M. (2011). Habit formation, naivete, and projection bias in gym attendance. *Working Paper.*

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*(9), 1082-1095.

Ariely, D., & Loewenstein, G. (2006). The heat of the moment: The effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making*, *19*(2), 87-98.

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science*, *13*(3), 219-224.

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes*, *37*(1), 93-110.

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *The Quarterly Journal of Economics*, *121*(2), 635-672.

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences, 30*(3), 241-254.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*(3), 610.

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*(6), 351-355.

Ben-David, I., Graham, J. R., & Harvey, C. R. (2010). *Managerial miscalibration* (No. w16215). National Bureau of Economic Research.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B.C., & Milkman, K. L. (2011). *The effect of providing peer information on retirement savings decisions* (No. w17345). National Bureau of Economic Research.

Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C. (2012). Active choice and health care costs: Evidence from prescription drug home delivery. *Working Paper*.

Bollinger, B., Leslie, P., & Sorensen, A. (2010). *Calorie posting in chain restaurants* (No. w15648). National Bureau of Economic Research.

Bond, S. D., Carlson, K. A., & Keeney, R. L. (2008). Generating objectives: Can decision makers articulate what they want? *Management Science, 54*(1), 56-70.

Bond, S. D., Carlson, K. A., & Keeney, R. L. (2010). Improving the generation of decision objectives. *Decision Analysis, 7*(3), 238-255.

Boorman, D. (2001). Today's electronic checklists reduce likelihood of crew errors and help prevent mishaps. *ICAO Journal*, *56*(1), 17-20.

Breman, A. (2011). Give more tomorrow: Two field experiments on altruism and intertemporal choice. *Journal of Public Economics*, *95*(11), 1349-1357.

Burson, K. A., Larrick, R. P., & Lynch, J. G. (2009). Six of one, half dozen of the other expanding and contracting numerical dimensions produces preference reversals. *Psychological Science, 20*(9), 1074-1078.

Cain, D. M., Loewenstein, G., & Moore, D. A. (2005). The dirt on coming clean: Perverse effects of disclosing conflicts of interest. *The Journal of Legal Studies*, *34*(1), 1-25.

Camerer, C. (1981). General conditions for the success of bootstrapping models. *Organizational Behavior and Human Performance*, *27*(3), 411-422.

Camerer, Colin F. (1998). Behavioral economics and nonrational decision making in organizations. In J. Halpern & B. Sutton (Eds.), *Decision Making in Organizations*. Ithaca, NY: Cornell University Press.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, *19*(1-3), 7-42.

Camilleri, A. R., & Larrick, R. P. (2013). Metric and Scale Design as Choice Architecture Tools. *Under Review.*

Carlson, K. A., Meloy, M. G., & Russo, J. E. (2006). Leader-driven primacy: Using attribute order to affect consumer choice. *Journal of Consumer Research, 32*(4), 513-518.

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.

Chapman, G. B., & Johnson, E. J. (2002). Incorporating the irrelevant: Anchors in judgments of belief and value. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge, UK: Cambridge University Press.

Chapman, G. B., Li, M., Colby, H., & Yoon, H. (2010). Opting in vs. opting out of influenza vaccination. *Journal of the American Medical Association*, *304*(1), 43-44.

Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica*, *77*(3), 909-931.

Cheema, A., & Soman, D. (2008). The effect of partitions on controlling consumption. *Journal of Marketing Research*, *45*(6), 665-675.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Erricsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 21-30). New York, NY: Cambridge University Press.

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in Experimental Social Psychology*, *24*(20), 1-243.

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, *108*(17), 6889-6892.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*(2), 95-106.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial Judgment. *Science, 243*(4899), 1668-1674.

Delaney, A. (2007). Eating behavior and portion control. Retreived from http://www.calorieking.com/library/article.php?content_id=3&path=13,66&art_id=843

Dunwoody, P. T. (2009). Theories of truth as assessment criteria in judgment and decision making. *Judgment and Decision Making, 4*(2), 116-125.

Einhorn, H. J., & Hogarth, R. M. (1975). Unit weighting schemes for decision making. *Organizational Behavior and Human Performance, 13*(2), 171-192.

Fagerlin, A., Wang, C., & Ubel, P. A. (2005). Reducing the influence of anecdotal reasoning on people's health care decisions: Is a picture worth a thousand statistics?. *Medical Decision Making, 25*(4), 398-405.

Federal Trade Commission. (1996). Protections for in-home purchases: The cooling-off rule. Retrieved from http://www.consumer.ftc.gov/articles/0176-protections-home-purchases-cooling-rule

Fennema, M. G., & Perkins, J. D. (2008). Mental budgeting versus marginal decision making: Training, experience and justification effects on decisions involving sunk costs. *Journal of Behavioral Decision Making, 21*(3), 225-239.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic and A. Tversky (Eds.) *Judgment Under Uncertainty: Heuristics and Biases* (pp. 422-444). Cambridge, UK: Cambridge University Press.

Fong, G. T., & Nisbett, R. E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology, 120*(1), 34-35.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives, 19*(4), 25-42.

Galesic, M., Garcia-Retamero, R., & Gigerenzer, G. (2009). Using icon arrays to communicate medical risks: Overcoming low numeracy. *Health Psychology, 28*(2), 210-216.

Gawande, A. (2010). *The checklist manifesto: How to get things right*. New York, NY: Metropolitan Books.

Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and vote turnout: Evidence from a large-scale field experiment. *American Political Science Review*, *102*(1), 33.

Gerber, A. S., & Rogers, T. (2009). Descriptive social norms and motivation to vote: Everybody's voting and so should you. *The Journal of Politics*, *71*(1), 178-191.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, *60*(4), 509-517.

Giné, X., Karlan, D., & Zinman, J. (2010). Put your money where your butt is: A commitment contract for smoking cessation. *American Economic Journal: Applied Economics*, 213-235.

Gino, F., Brooks, A. W., & Schweitzer, M. E. (2012). Anxiety, advice, and the ability to discern: Feeling anxious motivates individuals to seek and use advice. *Journal of Personality and Social Psychology, 102*(3), 497-512.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, *35*(3), 472-482.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology*, *38*(1), 69-119.

Hales, B. M., & Pronovost, P. J. (2006). The checklist: a tool for error management and performance improvement. *Journal of Critical Care*, *21*(3), 231-235.

Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. Oxford University Press on Demand.

Haran, U., Moore, D. A., & Morewedge, C. K. (2010). A simple remedy for overprecision in judgment. *Judgment and Decision Making, 5*(7), 467-476.

Harrahill, M., & Bartkus, E. (1990). Preparing the trauma patient for transfer. *Journal of emergency nursing: JEN: official publication of the Emergency Department Nurses Association*, *16*(1), 25-28.

Hart, E. M., & Owen, H. (2005). Errors and omissions in anesthesia: A pilot study using a pilot's checklist. *Anesthesia & Analgesia*, *101*(1), 246-250.

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, *31*(1), 405-440.

Heath, C., & Heath, D. (2010). *Switch: How to Change Things When Change is Hard*. New York, NY: Broadway Books.

Herzog, S. M., & Hertwig, R. (2009). The wisdom of many in one mind: Improving individual judgments with dialectical bootstrapping. *Psychological Science, 20*(2), 231-237.

Hogarth, R. M. (2001). *Educating Intuition*. Chicago, IL: University of Chicago Press.

Houser, D., Schunk, D., Winter, J., & Xiao, E. (2010). Temptation and commitment in the laboratory. *Institute for Empirical Research in Economics, University of Zurich Working Paper*, (488).

Huber, O., & Seiser, G. (2001). Accounting and convincing: The effect of two types of justification on the decision process. *Journal of Behavioral Decision Making*, *14*(1), 69-85.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing?. *Journal of Personality and Social Psychology*, *79*(6), 995-1006.

Jain, K., Mukherjee, K., Bearden, J. N., & Gaba, A. (In press). Unpacking the future: A nudge towards wider confidence intervals. *Journal of Consumer Research, 16*(2), 158–174.

Johnson, E. J., Bellman, S., & Lohse, G. L. (2002). Defaults, framing and privacy: Why opting in-opting out. *Marketing Letters*, *13*(1), 5-15.

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

Kahneman, D. (2011). *Thinking Fast and Slow*. New York, NY: Farrar, Straus and Giroux.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: a failure to disagree. *American Psychologist, 64*(6), 515.

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science, 39*(1), 17-31.

Kaur, S., Kremer, M., & Mullainathan, S. (2010). Self-control and the development of work arrangements. *The American Economic Review*, *100*(2), 624-628.

Keeney, R. L. (1992). *Value-Focused Thinking: A Path to Creative Decisionmaking*. Cambridge, MA: Harvard University Press.

Keeney, R. L. (2012). Value-focused brainstorming. *Decision Analysis, 9*(4), 303-313.

Kelley, T. (2001) The perfect brainstorm. In T. Kelley & J. Littman (Eds.), *The Art of Innovation: Lessons in Creativity from IDEO, America's Leading Design Firm* (pp. 53-66). New York, NY: Doubleday.

Keren, G. (1991). Calibration and probability judgments: Conceptual and methodological issues. *Acta Psychologica, 77*(3), 217-273.

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology, 55*(1), 623-655.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). *Rapid decision making on the fire ground.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society annual meeting.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 107-118.

Larrick, R. P. (2009). Broaden the decision frame to make effective decisions. In E. A. Locke (Ed.), *Handbook of Principles of Organizational Behavior* (pp. 461-480). Chichester, UK: Wiley and Sons.

Larrick, R. P., & Soll, J. B. (2008). The MPG illusion. *Science, 320*(5883), 1593-1594.

Larrick, R. P., Mannes, A. E., & Soll, J. B. (2012). The social psychology of the wisdom of crowds. In J. I. Krueger (Ed.), *Frontiers of Social Psychology: Social Psychology and Decision Making* (pp. 227-242). New York, NY: Psychology Press.

Larrick, R., Nisbett, R., & Morgan, J. (1993). Who uses the normative rules of choice. In R. E. Nisbett (Ed.) *Rules for Reasoning* (pp. 277-294). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lehman, D. R. & Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate education on reasoning. *Developmental Psychology, 26*(6), 952–960.

Lerner, J. S., Small, D. A., & Loewenstein, G. (2004). Heart strings and purse strings carryover effects of emotions on economic decisions. *Psychological Science*, *15*(5), 337-341.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*(2), 255.

Lewis, M. (2004). *Moneyball: The art of winning an unfair game*. New York, NY: W.W. Norton.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, *65*(3), 272-292.

Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Projection bias in predicting future utility. *The Quarterly Journal of Economics*, *118*(4), 1209-1248.

Lusardi, A., & Tufano, P. (2009). *Debt literacy, financial experiences, and overindebtedness* (No. w14808). National Bureau of Economic Research.

Ly, K., Mazar, N., Zhao, M., & Soman, D. (2013). A Practitioner's Guide to Nudging. *Research Report Series: Behavioral Economics in Action*. Rotman School of Management, University of Toronto. 15 March 2013.

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.

McCrea, S. M., Wieber, F., & Myers, A. L. (2012). Construal level mind-sets moderate self-and social stereotyping. *Journal of Personality and Social Psychology*, *102*(1), 51-68.

McKenzie, C. R., & Liersch, M. J. (2011). Misunderstanding savings growth: Implications for retirement savings behavior. *Journal of Marketing Research, 48*(SPL), S1-S13.

McKenzie, C. R., Liersch, M. J., & Finkelstein, S. R. (2006). Recommendations implicit in policy defaults. *Psychological Science*, *17*(5), 414-420.

Milkman, K. L. (2012). Unsure what the future will bring? You may overindulge: Uncertainty increases the appeal of wants over shoulds. *Organizational Behavior and Human Decision Processes*, *119*(2), 163-176.

Milkman, K. L., Akinola, M., & Chugh, D. (2012). Temporal distance and discrimination: An audit study in academia. *Psychological Science*, *23*(7), 710-717.

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2011). Using implementation intentions prompts to enhance influenza vaccination rates. *Proceedings of the National Academy of Sciences*, *108*(26), 10415-10420.

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2013). Planning prompts as a means of increasing preventive screening rates. *Preventive Medicine 56*(1), 92-93.

Milkman, K., Minson, J., & Volpp, K. (2012). Holding the hunger games hostage at the gym: An evaluation of temptation bundling. *The Wharton School Research Paper No. 45*. Available at SSRN: http://ssrn.com/abstract=2183859.

Milkman, K. L., Rogers, T., & Bazerman, M. H. (2008). Harnessing our inner angels and demons: What we have learned about want/should conflicts and how that knowledge can help us reduce short-sighted decision making. *Perspectives on Psychological Science*, *3*(4), 324-338.

Milkman, K. L., Rogers, T., & Bazerman, M. H. (2010). I'll have the ice cream soon and the vegetables later: A study of online grocery purchases and order lead time. *Marketing Letters*, *21*(1), 17-35.

Mitchell, D. J., Edward Russo, J., & Pennington, N. (1989). Back to the future: Temporal perspective in the explanation of events. *Journal of Behavioral Decision Making, 2*(1), 25-38.

Moore, D. A. & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502-517.

Myburgh, J. A., Chapman, M. J., Szekely, S. M., & Osborne, G. A. (2005). Crisis management during anaesthesia: Sepsis. *Quality and Safety in Health Care*, *14*(3), e22-e22.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: Theory and practice. *Annals of Behavioral Medicine, 35*(3), 261-274.

Nickerson, D. W., & Rogers, T. (2010). Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making. *Psychological Science*, *21*(2), 194-199.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, *89*(1), 103-124.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York, NY: Cambridge University Press.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17*(5), 407-413.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review, 111*, 781-799.

Read, D., & Van Leeuwen, B. (1998). Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Processes*, *76*(2), 189-205.

Rogers, T., & Bazerman, M. H. (2008). Future lock-in: Future implementation increases selection of 'should' choices. *Organizational Behavior and Human Decision Processes*, *106*(1), 1-20.

Royer, H., Stehr, M. F., & Sydnor, J. R. (2012). *Incentives, commitments and habit formation in exercise: Evidence from a field experiment with workers at a Fortune-500 company* (No. w18580). National Bureau of Economic Research.

Runciman, W. B., Kluger, M. T., Morris, R. W., Paix, A. D., Watterson, L. M., & Webb, R. K. (2005). Crisis management during anaesthesia: The development of an anaesthetic crisis management manual. *Quality and Safety in Health Care*, *14*(3), e1-e1.

Russo, J. E., & Schoemaker, P. J. H. (forthcoming). Overconfidence. Palgrave Encyclopedia of Strategic Management.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429-434.

Shah, A. K., Mullainathan, S., & Shafir, E. (2012). Some consequences of having too little. *Science, 338*(6107), 682-685.

Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, *26*(3), 278-292.

Simonson, I. (1989) Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16*(2), 158–174.

Simonson, I., & Nye, P. (1992). The effect of accountability on susceptibility to decision errors. *Organizational Behavior and Human Decision Processes*, *51*(3), 416-446.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research, 177*(3), 1333-1352.

Smith, S. M. (2003). The constraining effects of initial ideas. In P. B. Paulus & B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration* (pp. 15–31). Oxford, England: Oxford University Press.

Soll, J. B., & Klayman, J. (2004). Overconfidence in interval estimates. *Journal of Experimental Psychology-Learning Memory and Cognition, 30*(2), 299-314.

Soll, J. B., & Mannes, A. E. (2011). Judgmental aggregation strategies depend on whether the self is involved. *International Journal of Forecasting, 27*(1), 81-102.

Soll, J. B., Keeney, R. L., & Larrick, R. P. (2011). Consumer misunderstanding of credit card use, payments, and debt: Causes and solutions. *Journal of Public Policy & Marketing*, (ja), 1-49.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate?. *Behavioral and Brain Sciences, 23*(5), 645-665.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94*(4), 672-695.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology, 73*(3), 437-446.

Thaler, R. H. (June 2012). Smart disclosure: The regulatory future. Boulder Summer Conference on Consumer Financial Decision Making, Boulder, Colorado.

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*. New Haven, CT: Yale University Press.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*(7), 1275-1289.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*(3), 403-421.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440-463.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Ubel, P. (2012). Beyond comprehension: Figuring out whether decision aids improve people's decisions. In E. Shafir (Ed.), *The Behavioral Foundations of Public Policy* (pp. 351-361). Princeton, NJ: Princeton University Press.

United States Air Force Series:—OV-10A Bronco Association. Pilots' Abbreviated Flight Crew Checklist—USAF Series: OV-10A Aircraft. October 1999 (March 2006):1.

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive–based approaches for weight loss. *Journal of the American Medical Association*, *300*(22), 2631-2637.

Volpp, K. G., Levy, A. G., Asch, D. A., Berlin, J. A., Murphy, J. J., Gomez, A., ... & Lerman, C. (2006). A randomized controlled trial of financial incentives for smoking cessation. *Cancer Epidemiology Biomarkers & Prevention*, *15*(1), 12-18.

Volpp, K. G., Loewenstein, G., Troxel, A. B., Doshi, J., Price, M., Laskin, M., & Kimmel, S. E. (2008). A test of financial incentives to improve warfarin adherence. *BMC Health Services Research*, *8*(1), 272.

Vul, E., & Pashler, H. (2008). Measuring the crowd within probabilistic representations within individuals. *Psychological Science, 19*(7), 645-647.

Wansink, B., & Cheney, M. M. (2005). Super bowls: Serving bowl size and food consumption. *Journal of the American Medical Association*, *293*(14), 1727-1728.

Wolff, A. M., Taylor, S. A., & McCabe, J. F. (2004). Using checklists and reminders in clinical pathways to improve hospital inpatient care. *Medical Journal of Australia*, *181*, 428-431.

Yaniv, I., & Milyavsky, M. (2007). Using advice from multiple sources to revise and improve judgments. *Organizational Behavior and Human Decision Processes, 103*(1), 104-120.