A USER'S GUIDE TO DEBIASING

Jack B. Soll
Fuqua School of Business
Duke University

Katherine L. Milkman
The Wharton School
The University of Pennsylvania

John W. Payne
Fuqua School of Business
Duke University

Forthcoming in *Wiley-Blackwell Handbook of Judgment and Decision Making*
Gideon Keren and George Wu (Editors)

Improving the human capacity to decide represents one of the great global challenges for the future, along with addressing problems such as climate change, the lack of clean water, and conflict between nations. So says the Millenium Project (Glenn, Gordon, & Florescu, 2012), a joint effort initiated by several esteemed organizations including the United Nations and the Smithsonian Institution. Of course, decision making is not a new challenge—people have been making decisions since, well, the beginning of the species. Why focus greater attention on decision making now? Among other factors such as increased interdependency, the Millenium Project emphasizes the proliferation of choices available to people. Many decisions, ranging from personal finance to health care to starting a business, are more complex than they used to be. Along with more choices comes greater uncertainty and greater demand on cognitive resources. The cost of being ill-equipped to choose, as an individual, is greater now than ever.

What can be done to improve the capacity to decide? We believe that judgment and decision making researchers have produced many insights that can help answer this question. Decades of research in our field have yielded an array of *debiasing* strategies that can improve judgments and decisions across a wide range of settings in fields such as business, medicine, and policy. And, of course, debiasing strategies can improve our personal decisions as well. The purpose of this chapter is to provide a guide to these strategies. It is our hope that the ideas in this chapter can immediately be applied, so that 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.

What is debiasing?

Before proceeding further, it is important to define what we mean by "debiasing". We consider a bias to be a deviation from an objective standard, such as a normative model (see Baron, 2012). For example, according to the economic view of rationality, decisions should be based on beliefs about possible outcomes, their associated values or utilities, and their probabilities of occurrence. Yet research on judgment and decision making has demonstrated

numerous violations of this principle, such as preference reversals, framing effects, and the inappropriate weighting of extreme probabilities (e.g., see chapters A, B, and C in this volume). Similarly, the normative model of discounting does not allow for systematic intertemporal preference reversals (e.g., preferring $25 in 51 weeks to $20 in 50 weeks, but preferring $20 today to $25 in 1 week; Prelec & Lowenstein, 1991). Thus, we would consider a person who repeatedly *plans* to eat healthily yet consistently gives in to tempting snacks to be worthy of debiasing. Note that we may also want to help the person who plans to eat unhealthily and does so, with little regard for future health consequences or the resulting burden on the health care system, but this is not an example of debiasing and therefore not a subject of this chapter.

Our treatment of debiasing includes addressing both coherence-based biases that reflect logical inconsistencies (e.g., as defined by probability theory or economics), and correspondence-based biases that reflect systematic misperceptions or misjudgments of reality (Hammond, 1996). Further, in some cases, inaccurate judgments themselves may not be systematically biased, but the process that produces them is systematically deficient in some way. For example, in forming judgments people tend to use available information both inconsistently and incompletely, and this can detract from accuracy. We consider techniques that improve judgment by addressing these deficiencies to be examples of debiasing as well.

A second distinction can be made between debiasing and the broader topic of improving decisions. One way to improve decisions is to provide new information (e.g., telling people about some new available options). This is not debiasing because people may be doing the best they can with what they know. However, sometimes existing information can be reframed in a way that highlights its importance or corrects a misunderstanding, and we do call this debiasing. For example, American retirees can choose to start receiving social security benefits anytime between the ages of 62 and 70. By delaying until age 70, a retiree can secure larger payments that help insure against the prospect of outliving her money. Yet many people opt for the much smaller payments that begin at age 62. Clearly, not everyone should delay; some people may need the money or expect to die relatively young. One way to potentially improve this decision

would be to calculate and graphically present the time-path of financial resources a retiree would have available given different choices about when to start receiving payments. This recalculation could be considered new information, especially for those who cannot do the math on their own. However, we consider it to be a type of debiasing, rather than a form of new information, because it helps people make better use of the information already available to them. With this in mind, we see debiasing as a continuum, ranging from the reframing or repackaging of existing information, to the provision of new strategies for thinking about information.

Types of Debiasing

Our categorization of debiasing methods builds on Fischhoff's (1982) classic distinction that attributes biases to either persons or tasks. When attributing bias to the person, one implicitly assumes that the situation is more or less fixed, and therefore the best approach is to provide people with some combination of training, knowledge, and tools to help overcome their limitations and dispositions. We dub this approach "modify the decision maker." It 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 (Arkes, 1991; Larrick, 2004). For example, people often delay saving for retirement, partly due to the mistaken belief that investments grow linearly over time (Stango & Zinman, 2009). Because, other things being equal, savings at a constant rate of interest actually grow exponentially, people who start saving early in their careers will be dramatically better prepared. To combat the faulty thinking of those who believe investments grow linearly, people can be taught about compound interest, or taught simple approximations such as the "rule of 72" (if X is the annual interest rate, money doubles approximately every 72/X years).

The second approach, which we call "modify the environment," seeks to alter the environment to provide a better match for the thinking that people naturally do when unaided (Klayman & Brown, 1993), or alternatively, to encourage better thinking. We pause here, because these are two very different ways to modify the environment. One general approach is to

change something about the situation that spurs people to process information more appropriately. For example, when considering retirement savings options, employees could be shown graphs displaying how wealth would grow over time under different scenarios for annual contributions (McKenzie & Liersch, 2011). A second approach adapts the environment to people's biases. In the case of savings, this idea is illustrated by Thaler and Benartizi's (2004) popular and effective Save More Tomorrow™ plan, which encourages employees to increasing their contributions, but only out of future raises. This allows savers to sidestep loss aversion (since current spending is not reduced), and takes advantage of choosing in advance, a debiasing method we describe later in this chapter. Save More Tomorrow™ is an example of a *nudge*—an intervention that modifies the environment without restricting choice or altering incentives in a significant way (Thaler & Sunstein, 2008). Nudges rely on psychological principles to influence behavior for the good of the individual or society (as opposed to for the good of the nudger, in which case they would be indistinguishable from many marketing tactics). When used judiciously, nudges can be very helpful for debiasing the individual, which is our focus in this chapter.

Our discussion of retirement savings also highlights another distinction. A given debiasing method may be geared toward producing a specific outcome (e.g., everyone saves more), or an improved process that could lead to variety of outcomes (e.g., everyone saves the right amount for themselves). We believe that both types of methods are useful. Some situations call for a blunt instrument that nudges everyone in the same direction, whereas others (when individuals are heterogeneous in their preferences) require a more refined approach that helps people make better decisions for their own unique circumstances (Dietvorst, Milkman and Soll, 2014; Fernandes, Lynch, & Netemeyer, 2013; Johnson, Hassin, Baker, Bajger, & Treuer, 2013).

Chapter Overview

We begin 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 refers to whether an individual is in a position to make a good decision in a particular situation. Intense emotional states, fatigue, and poor decision-related skills (e.g., being innumerate) can all contribute to a lack of decision readiness. We then turn to a review of debiasing techniques, organized according to whether they modify the person or the environment. We close with a discussion of six considerations in choosing which debiasing method to apply.

SOURCES OF BIAS

System 1 and System 2

Whether one chooses to modify the person or the environment, it often helps to know the psychological factors that contributed to produce a bias in the first place. Although a complete theory would likely model the human mind as comprised of multiple, interrelated processes, many scholars have found it useful to conceive of decision making as guided by two mental systems—*System 1* and *System 2* (Kahneman, 2011; for a critique of this distinction, see Keren & Schul, 2009). *System 1* refers to processes that are fast, effortless, and automatic, like memory or a fight-or-flight response. Although System 1 often gets things right, speed and efficiency come with a cost: systematic errors when we encounter certain types of judgment problems (Arkes, 1992). For example, imagine a relative approaches you with an investment opportunity in e-commerce. The idea evokes favorable comparisons with several highly successful companies (e.g., started in a garage by a quirky but brilliant Ivy League dropout), which in turn fuels an intuitive, System 1 sense of optimism about the project. This enthusiastic first impression is likely to have ignored two important facts. First, similar projects that failed are unlikely to come to mind (because you never heard of them in the first place)—an example of what Kahneman (2011) calls WYSIATI (*What You See Is All There Is*). System 1 judgments are based on information retrieved from memory and perceived by the senses; the fact that what you retrieve is skewed toward certain kinds of information (examples of success) and not toward other kinds (examples of failure) is typically not taken into account. Second, the base rate of

success for e-commerce ventures is quite low. Most new businesses fail, and your relative's project is more likely to wind up among the ones that fizzle out than among the stock market darlings.

This is where *System 2* comes in—the slower, more deliberate, more conscious kind of thinking involved in paying close attention to a lecture or solving a complicated math problem. To correct a faulty intuition, one needs to pay attention and reflect at least briefly before arriving at a final judgment (Kahneman & Frederick, 2005). But this is not enough. One also needs at least a rough understanding of the correct normative rule and needs to apply it reasonably well (Larrick, 2004). For example, in the aforementioned investment example, knowledge of Bayes' rule could help in thinking through how a low success base rate might be revised in light of the evidence about the project. According to Bayes' rule, the odds of success are obtained by multiplying the prior odds (i.e., the base rate) by the likelihood ratio, which in this case captures the relative chances of observing the investment situation you face among successful as opposed to unsuccessful companies. Our example is deliberately vague so that it is similar to the fuzziness of real-world problems. Even so, thinking through the Bayesian logic should dramatically temper optimism if one realizes that many ventures started in a garage never make it out the door. But System 2 cannot help if knowledge or understanding of the relevant normative principle is lacking.

One can also know the normative rule but still not apply it. Depending on the circumstances, this could be characterized as a System 2 failure of attention and monitoring, cognitive laziness, or an adaptive adjustment to a challenging or unrewarding environment (Payne, Bettman, & Johnson, 1993). Also, even a person who never learned Bayes' Rule could still potentially reflect upon the problem and think through some of the logic, such as recognizing that the base rate is a good anchor for judgment, and recognizing that apparent diagnostic evidence may be exaggerated as System 1 selectively attends to facts and arranges them into a cohesive story that makes sense (Kahneman, 2011).

Narrow Thinking

The entrepreneurial investment decision highlights a common problem. People often form quick and intuitive judgments based on limited information—either retrieved from memory or delivered by the environment—which may be incomplete, ambiguous, or biased (Hogarth, 2001). However, even intense deliberation can produce narrow thinking if it does not occur to the decision maker that something is missing or that there are alternative perspectives.

To illustrate, consider the following study by Bond, Carlson, and Keeney (2008). The researchers asked masters of business administration students to list all of the objectives relevant to choosing a summer internship after their first year in the program. When shown a master list of objectives generated by others (which included items such as "improves my attractiveness for full-time job offers," and "helps me develop my leadership skills"), the average student checked off fifteen objectives, of which only about half had been previously self-generated. They saw value in the objectives that their classmates had come up with, yet without guidance they would have likely made life-changing decisions without taking these into consideration. 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 (essentially, a System 2 failure to monitor and intervene). About half of 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. As illustrated by the Gestalt psychologist Duncker (1945), when solving a problem people often become fixated on one type of solution, which impedes the ability to generate other, more creative ideas. Additional generated ideas are thus often variants on a theme rather than truly novel (Smith, 2003). Narrow thinking also contributes toward many classic judgment and decision making biases. To take just one example, consider the anchoring bias. When estimating a numerical answer such as how long it will take the Dow Jones Industrial Average to reach 20,000, people are typically overly influenced by a starting value. The 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 20,000?"), another person's opinion, or even an arbitrary number generated at random (Tversky & Kahneman, 1974).

Anchors are influential partly 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). Relatedly, when making choices people tend to interpret ambiguous information in a manner that favors the option toward which they are tentatively leaning (Carlson, Meloy, & Russo, 2006). Together, this research suggests a valuable approach to debiasing—broaden thinking to consider disconfirming evidence and alternative interpretations (Larrick, 2009).

DECISION READINESS

Deliberate, System 2 style thinking plays a critical role in decision making—it monitors intuitive judgment and, when necessary, corrects it (Kahneman & Frederick, 2005). This is true even for experts, for whom strong intuitive associations based on contextual cues can occasionally lead to glossing over details that are atypical but important (Chi, 2006). We call a person "decision ready" when System 2 is capable of performing its functions, in terms of monitoring, suspending decisions, and correcting judgments. What determines whether a person is decision ready or not? We highlight three factors.

- *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 ability to monitor decisions and notice possible errors. Tasks that require effort and attention such as exercising self-control become more difficult (Baumeister et al., 2007; Milkman, 2012). Distraction and time pressure have a similar effect, redirecting attention toward finding quick solutions and greater reliance on heuristic, System 1 processing (e.g., Gilbert & Hixon, 1991; Payne et al., 1993).
- *Visceral influences.* Visceral reactions are essential to survival; recoiling in fear from a snake before reflection is a wise move. Even so, emotions and other visceral experiences

do sometimes negatively impact decisions. People behave as if present desires will remain the same in different physical or emotional states (Loewenstein, 1996). This empathy gap can lead hungry shoppers, for instance, to 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). 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). For example, compared to a control group, participants who watched a video clip that made them feel sad were more likely to prefer a smaller, sooner reward than a larger, later one (Lerner, Li, & Weber, 2013). A follow-up study showed that gratitude has the reverse effect, inducing greater patience (DeSteno, Li, Dickens, & Lerner, 2014).

- *Individual Differences.* People differ in 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). Also, some people are more reflective and are therefore more likely to detect situations in which careful reasoning is needed (Toplak, West, & Stanovich, 2011).

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. One way to improve decision making is to simply avoid making important decisions when depleted, fatigued, angry, aroused, hungry, distracted, or untrained. But can people recognize their own unreadiness? 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 successfully detect those same biases 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.

Modifying the environment offers an alternative route to improving decision readiness. Consider the case of Israeli judges making a series of parole decisions (Danziger, Levav, & Avnaim-Pesso, 2011). The judges were relatively lenient immediately following meal breaks, but as time elapsed following each break they denied parole requests with greater frequency, more often sticking with the default decision that keeps the applicant in prison. This inconsistency and arbitrariness could arguably be cured (at least partially) with a simple environmental modification—scheduling more breaks to nudge judges toward readiness. Environmental modifications can also be used to 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.

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. Students with coursework in economics or statistics are more likely to successfully apply fundamental principles from those disciplines to avoid biases (see review in Larrick, 2004). People can also learn to apply statistical principles in a lab context and extend the learning to new domains, although transfer tends to fade over a span of weeks (Fong & Nisbett, 1991). Relatedly, economics professors, as opposed to those in the humanities or biology, are more likely to apply economic principles in life, such as ignoring sunk costs (Larrick, Morgan, & Nisbett, 1990).

Education makes a difference, but how much? If individuals are to debias themselves, they must not only possess the correct normative principles, but also identify the situations in which

to apply those principles and be motivated to do so. Training in normative rules often fails when people have strong intuitions and do not pause to think more deeply (McKenzie and Liersch, 2011). The most effective type of education is domain-specific training on a decision task that will be engaged in very soon and possibly repeatedly. For example, probability judgments are typically miscalibrated—when judges are 90% sure that an event will occur, for instance, the actual rate of occurrence is typically far lower. But experts such as meteorologists are surprisingly well-calibrated when predicting within their domain of expertise, provided that they receive training, timely feedback, and appropriate incentives. 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 poorly calibrated as the rest of us (Keren, 1991).

This is not 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 many other biases (Peters et al., 2006). Environmental modifications, such as providing transparent disclosure in financial statements, can alleviate some of the costs of being innumerate (Soll, Keeney, & Larrick, 2013). Nevertheless, basic quantitative skills are important. An open question for debiasing research is to determine not only whether people can retain learned skills, but also whether they can reliably apply their skills when the situation calls for it (Fernandes et al., 2014).

Cognitive Strategies

We all know at some level that it helps to look at problems from multiple perspectives. Yet people frequently fail to do so 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. In this section, we present several examples of how rules such as “think of the opposite” and “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. 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. For example, Keeney lists 29 objectives for selecting a title for a journal article (e.g., communicative of content, understandable, clever, etc.). By focusing on these 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.

Tempering Optimism

On any new undertaking, people often overestimate their chances of success (Moore & Healy, 2008; see also chapter D). This results from a tendency to focus too narrowly on evidence that supports an initial (often preferred) hypothesis and to underweight contradictory evidence. One fix is to simply "think of the opposite" by articulating reasons why an initial answer might be wrong, or why a project idea might fail (Koriat, Lichtenstein, & Fischhoff, 1980). Russo and Schoemaker (2013) 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 2015 might ask "Here in 2035, 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 dampen excessive optimism and therefore spur decision makers to plan for a range of possible contingencies.

Improving Judgmental Accuracy

When judgments are provided by many people, an extremely effective way to combine them is to weight them equally, such as by taking the simple average or applying majority rule (e.g., Clemen, 1989; Hastie & Kameda, 2005). The idea of harnessing the "wisdom of crowds" has been applied to a wide variety of contexts, ranging from sports prediction markets to national security (Surowiecki, 2004). For quantity estimates, averaging provides benefits over the average individual whenever individual guesses bracket the truth (i.e., some guesses on both sides), so that high and low errors will cancel out (Larrick, Mannes, & Soll, 2012). Remarkably, the same method can be applied when there is only one person by taking advantage of "the crowd within." The underlying insight is that on any given judgment people use only a subset of the accessible information (notice the similarity to Kahneman's WYSIATI). By answering the same question twice, a person might retrieve from memory somewhat different samples of evidence and provide different answers. Typically, averaging these within-person answers provides about half of the accuracy gain that could be achieved by averaging guesses from two different people (Larrick et al., 2012).

A simple way to harness the crowd within is to introduce a time delay between two answers (Vul & Pashler, 2008). For example, a corporate analyst might predict sales for a set of products, and then go through the pile again to produce another batch of forecasts to be averaged with the first. This method could help compensate for low decision readiness due to fatigue or time pressure because nonsystematic errors produced by these factors will tend to cancel out when estimates are averaged. An alternative procedure, Herzog and Hertwig's (2014) dialectical bootstrapping, dispenses with delay. After making an initial guess, the judge follows up by assuming the guess is wrong and makes a second guess. Yet another method asks the judge to generate multiple estimates using different thinking styles or strategies, such as a quick System 1 guess followed by a more deliberate System 2 response (Larrick & Soll, 2013). All of these techniques encourage people to consider new evidence that might support different answers. Averaging the two judgments tends to outperform trying to identify the better of the two, partly

because answers based on different pools of evidence often bracket the truth, and partly because people are imperfect at guessing which answer is better.

Assessing Uncertainty

One of the most robust forms of overconfidence arises on interval judgments (e.g., "I am 80 percent sure that the house will sell for between 250 and 275 thousand dollars"), where wider intervals indicate greater uncertainty. For example, over a nine-year time horizon, Ben-David, Graham, and Harvey (2013) 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 outcomes.

Although it is difficult to completely erase this type of overconfidence, three 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 a range estimate. The improved performance arises from intervals that are both wider and better centered on the truth (Soll & Klayman, 2004). A second method provides judges with a series of small-interval bins (e.g., 0-10, 11-20, etc.); the judge assigns a probability to each bin and thereby maps out a probability distribution (Haran, Moore, & Morewedge, 2010). The aforementioned methods work because they encourage people to consider evidence for a broad spectrum of answers, as opposed to just for their best guess (again, a cure for narrow thinking). The final method applies to forecasting time series. Rather than forecasting the price of 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. 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 (Jain, Mukherjee, Bearden, & Gaba, 2013).

Use Models to Decide

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 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 judgment alone) acquired a performance advantage until their competitors began to also develop and apply equations.

Linear models outperform expert judgment across a wide range of settings (Dawes, Faust, & Meehl, 1989). The most sophisticated approach 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 to 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, 1979). 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 *bootstrap* models outperform the very expert judgments they model by reducing the noise inherent in experts' decision rules (Camerer, 1981).

Linear models systematize the reliance on relevant decision criteria and eliminate the opportunity for bias to creep into a decision and reduce its quality. When human judgment is critical to predicting outcomes (e.g., a rating of an applicant's essay), it can be recorded numerically and entered as an input into a model. In fact, a linear model that sums a set of subjective ratings can be highly predictive, such as the APGAR test for assessing a newborn infant's health (Casey, McIntire, & Leveno, 2001). However, models do have some limitations. First, they are only as valuable as the attributes they include. If important inputs are overlooked (e.g., due to narrow thinking), the solution produced may be error-prone or biased. Second, under special or changing circumstances, models based on historical data will not apply and may even lead the decision maker wildly astray. Determining when and where models do not apply is one of the greatest challenges associated with relying on these valuable decision tools.

Linear models are just one of many types of quantitative models that can help systematize judgments in order to reduce opportunities for error. Related to models are checklists, which improve the consistency of repeated decisions. We will discuss checklists as a debiasing tool later in this chapter. See chapter E in this volume for a discussion of other useful quantitative modeling 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 bias. We will begin with a discussion of incentives; perhaps people would make wiser choices if there were a greater payoff for doing so. We then turn our attention to modifying the environment by a different means—using well-understood psychological principles as tools to improve 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 retirement savings plan) to "*nudge*" choices in wise directions (Thaler & Sunstein, 2008). Although much of the work on choice architecture assumes that a policy maker shapes the environment, people can also sometimes modify the environment

for their own future selves—they can be their own choice architects. Finally, we will examine how organizations might alter the work environment through cognitive repairs in ways that encourage better decisions (Heath, Larrick, & Klayman, 1998). This kind of debiasing often arises bottom-up from employees themselves, and tends to have a greater social element than other forms of modifying the environment.

Incentives

It is sometimes suggested that decision biases arise due to insufficient motivation, and therefore one way to modify the environment is to pay people to make smarter choices. One way incentives might work is to give people a reason to shift from mindless, System 1 thinking to more mindful, System 2 deliberation. Decades of research into the effectiveness of incentives paint a very mixed picture (see review by Camerer & Hogarth, 1999). Incentives do often help. When it comes to self-control problems, incentives can motivate people to resist temptation, and thus can help them lose weight (Volpp et al., 2008), quit smoking (Volpp et al., 2006), and exercise more (Charness & Gneezy, 2009; Acland & Levy, 2013). In settings where habit formation is critical to long-term behavior change, incentives provide an incipient motive to engage in the desired behavior. The beauty of it is that the incentive can often be removed once the desired habit is formed, or the undesired habit eradicated (Charness & Gneezy, 2009; Acland & Levy, 2013).

Incentives have a much weaker effect for biases that are not primarily caused by lack of effort or insufficient attention. Monetary incentives can even backfire in some instances, by leading people to "think harder but not smarter" by investing more cognitive effort into incorrect models and theories (Larrick, 2004). In one notable example, incentivized individuals were more likely to try to improve upon reliable formulas by applying idiosyncratic knowledge, which hampered their performance compared to others who were not rewarded for accuracy (Arkes, Dawes, & Christensen, 1986). Relatedly, when there is a large chance component in the environment (e.g., the stock market), incentives can cause people to look for and find patterns

that are not there, and consequently perform worse than a default or "base-rate" policy based on long-run historical trends and averages.

Increased accountability is another type of incentive—it increases the cost of failure and thus people's motivation to make good decisions (see review by Lerner & Tetlock, 1999). When people expect that they will have to explain their decisions to others, they may invest more effort in solving a problem, and are more likely to arrive at a normative solution provided that they know the correct rule and how to apply it. Thus, accountability has been shown to help people ignore sunk costs (Simonson & Nye, 1992), as one of many examples. Just like monetary incentives, however, accountability is unlikely to help when cognitive laziness is not the root source of bias.

Choice Architecture

As described earlier, choice architecture refers to the manner in which alternatives or information is presented to decision makers, and a choice architect is someone who is actively engaged in designing that environment (Johnson et al., 2013; Thaler & Sunstein, 2008). For a modification of the environment 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). An example can help illustrate this concept: Making enrollment in a company's retirement savings program the default is a nudge that can increase enrollment rates by as much as 37 percentage points (Madrian & Shea, 2001). Although defaults do not restrict choice (people can "opt-out"), people nevertheless exhibit inertia in the face of a default (Johnson, Bellman, & Lohse, 2002). In the retirement savings example, this inertia is leveraged to help people overcome the common decision-making pitfall of present-bias—the tendency to overly discount the future. 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

There are many settings where defaults have been used successfully to achieve the aims of policy makers seeking to improve individual and social welfare. For example, 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). In part, defaults work because they leverage decision makers' inertia. Those who procrastinate, are preoccupied, or are otherwise oblivious are automatically opted in. Defaults can therefore help individuals who repeatedly say they will "do it later," for instance, when it comes to making a flu shot appointment (Chapman, Li, Colby, & Yoon, 2010). There are other mechanisms as well—defaults may establish a reference point that loss aversion makes painful to sacrifice (Johnson et al., 2002), and they may implicitly suggest a recommended course of action that is perceived as expert advice (McKenzie, Liersch, & Finkelstein, 2006).

The fact that default effects have many causes contributes to their power and robustness, but it is also a source of criticism because there are many ways in which a default could leave an individual with an outcome ill-suited to their personal preferences. The one-size-fits-all nature of many defaults, which do not typically account for preference heterogeneity, may benefit many people at the expense of harming a few (for a thorough and excellent discussion of the benefits and limitations of defaults, see Smith, Goldstein, & Johnson, 2013). Despite their potential drawbacks, we believe that defaults have a place in the debiaser's toolkit, especially when they can be crafted to fit the context, for instance through the use of algorithmically-generated "smart defaults" that infer the optimal choice for an individual from his or her demographic traits and other personal information (Smith et al., 2013).

Nudges that 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 upon their objectives and plans. Soman, Xu, and Cheema (2010) call this moment when people could improve decisions through elaboration a "decision point," and define

it as the time when people can potentially avoid errors due to mindlessness or visceral impulses. Below, we discuss several types of interventions that prompt additional thought.

Planning Prompts. Prompting the formation and articulation of concrete plans to complete a desired action can help decision makers avoid follow-through failures due to both procrastination and forgetfulness. When an individual is prompted to decide on the when, where, and how of an intended action, the newly formed plan represents a commitment that is both psychologically difficult to break and memorable (Gollwitzer & Sheeran, 2006). Planning prompts have been used successfully to increase many important outcomes including exercise (Milne, Orbell, & Sheeran, 2002), meeting deadlines (Dholakia & Bagozzi, 2003), and flu shot take-up (Milkman, Beshears, Choi, Laibson, & Madrian, 2011).

Planned Interruptions. Another way to reduce judgment errors resulting from an under-reliance on System 2 reasoning is to build interruptions into choice environments in order to encourage added reflection. For example, interrupting the consumption of food (e.g., by partitioning the same quantity into several separate containers rather than one container) can reduce mindless eating (Cheema & Soman, 2008), and partitioning cash wages into multiple envelopes can increase saving (Soman & Cheema, 2011). By partitioning resources, choice processes are interrupted at the moment when a partition is encountered, and these interruptions lead to slower decisions and deeper processing. Such interruptions are most valuable when imposed on decisions where an individual intends to regulate consumption but sometimes fails due to mindlessness.

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. 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) increased 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 increased 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.

Nudges that 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 encouraging less impulsive, more reasoned decisions is to prompt individuals to decide well in advance of the moment when those decisions will take effect. Choosing in advance has been shown to increase people's support for "should" decisions, or those that provide delayed benefits but short-term pain (e.g., saving more for retirement, exercising, eating healthily) (Rogers & Bazerman, 2008; Milkman, Rogers and Bazerman, 2010). Another result of choosing in advance is that people make choices in a higher construal-level mindset, which means they focus more on abstract objectives (e.g., why?) rather than concrete plans (e.g., how?) (Trope & Liberman, 2003). Thus, choosing in advance has the potential to facilitate greater consideration of one's objectives in making a decision.[1]

Pre-commitment. People tend to make more patient and reasoned decisions for the future than for the present. Therefore, providing opportunities for individuals to *both* choose in advance *and* make a binding decision (or at least a decision where penalties will accompany a reversal)

[1] A byproduct of this, however, is that 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 for certain types of decisions, as demonstrated in a field study of decisions on whether to grant prospective graduate students requests for meetings (Milkman, Akinola, & Chugh, 2012).

can improve many choices (Ariely & Wertenbroch, 2002). For example, people save substantially more with bank accounts that have 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. In a study of Philippine bank customers, Ashran, Karlan, and Yin (2006) found that individuals who were offered a choice between a commitment account and an unconstrained account with the same interest rate saved, on average, 81% more than those in a control group who had access only to the unconstrained account. 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 reduces consumption substantially (Wansink & Cheney, 2005).

Temptation Bundling. A new twist on pre-commitment called "temptation bundling" 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). The decision maker commits to engaging in the gratifying, indulgent activity only when simultaneously engaged in the virtuous activity. The result: increased engagement in beneficial behaviors like exercise and reduced engagement in guilt-inducing, indulgent behaviors (Milkman, Minson, & Volpp, 2014).

Nudges that Kindly Shape Information

People are more likely to reach accurate conclusions when they have the right information packaged in an intuitively comprehensible and compelling 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 they fail to detect the necessity). For example, consumers spend less when unit pricing information (e.g., the price per ounce of a product) is displayed not only on each product tag individually, but also on

an organized list that makes it even easier for consumers to compare prices (Russo, 1977). In the parlance of Hsee (1996), the organized list makes price more evaluable, shifting weight to that attribute. Below we provide examples of several additional strategies that can be used to shape and package information so it will be particularly impactful for the purposes of debiasing.

Transform the scale. Metrics such as MPG (miles per gallon) for vehicles, SEER (seasonal energy efficiency ratio) ratings for air conditioners and megabytes per second for data transfer share a common property—the relationship with the variable relevant to the consumer's objective (e.g., minimizing fuel consumption, time) is nonlinear. For example, a change in MPG from 10 to 11 saves just as much gas as a shift from 33 to 50 (1 gallon per 100 miles), but the latter is perceived as having a much greater impact. Research by Larrick and Soll (2008) showed that (1) improvements at the low end of MPG (e.g., introducing hybrid trucks) tend to be undervalued; and (2) providing consumers with GPhM (gallons per hundred miles) leads to more accurate perceptions because GPhM is linearly related to consumption and cost. As a consequence of this research, GPhM is now included on federally mandated US vehicle labels.

Expand the scale. The new federally-mandated vehicle labels also state fuel-cost savings over 5 years compared to an average new vehicle. This metric could have been provided on a different scale (e.g., 1 month, 1 year, etc.), but arguably the 5-year time frame is appropriate because it matches the typical vehicle ownership period and places gas consumption in the context of other large purchases. Similarly, people weight fuel costs more heavily when expressed in terms of the lifetime miles traveled (e.g., $17,500 per 100,000 miles rather than a smaller scale; Camilleri & Larrick, 2014). The underlying principle here is that, within reason, larger scaling factors cause people to weight an attribute more heavily (Burson, Larrick, & Lynch, 2009).

Frame messages appropriately. When providing information for a decision, the communicator often has the option of framing outcomes in terms of either gains or losses. Since the introduction of prospect theory (Kahneman & Tversky, 1979), scholars have explored the subtle ways in which frames shift reference points, and the implications for decision making.

Framing effects are often dramatic, and thus the framing of persuasive messages has great potential as a debiasing tool. Consider, for example, Rothman & Salovey's (1997) application of prospect theory principles to messaging in the health domain. As they predicted, loss-framed messages are typically superior for promoting illness detection behaviors, and gain-framed messages are superior for promoting illness prevention behaviors (see review and discussion of mechanisms by Rothman & Updegraff, 2010). The pattern suggests, for example, that a message designed to promote screening for colon cancer should focus on averting potential losses (e.g., "helps avoid cancer" as opposed to "helps maintain a healthy colon"), whereas a message to promote regular exercise should focus on reaping the gains (e.g., "increases life expectancy" as opposed to "lessens risk of heart disease").

Use kind representations for guidelines. 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 (e.g., fruits, vegetables, grains, etc.). The guide was too abstract to be useful (Heath & Heath, 2010). The USDA's new MyPlate diagram provides a more intuitive model, showing a picture of a plate ideally divided across the food groups. Half the plate is filled with fruits and vegetables.

Use kind representations for probabilities. Probabilistic information is notoriously confusing, and providing relative frequency information (e.g., 1 out of every 10,000 instead of 0.01%) can help (Hoffrage et al., 2000). Ideally, new representations lead decision makers to better understand the deep structure of the problem they face (Barbey & Sloman, 2007). One promising method for conveying probabilistic information is through visual 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% chance of success, or a less arduous procedure, angioplasty, with a 50% chance of success. Participants relied much less on irrelevant anecdotal information in making decisions when the procedures' stated success probabilities were accompanied by 10x10 grids of differently colored or shaded icons to visually represent the relative frequencies of success and failure.

Convey social norms. Individuals have a tendency to herd, or to imitate the typically observed or described behaviors of others (Cialdini, Kallgren, & Reno, 1991), in part because the behavior of the herd often conveys information about wise courses of action but also in part due to concerns about social acceptance. This tendency can be used strategically: Providing information about the energy usage of one's neighbors on an electricity bill (rather than only conveying information about one's own usage) can reduce energy consumption by 2% (Alcott, 2011). Providing social norms can sometimes backfire—the strategy is most effective when the desired outcome is seen as both popular and achievable. For example it can be demotivating to learn that the majority of others are so far ahead on retirement savings that it will be hard to catch up (Beshears, Choi, Laibson, Madrian, & Milkman, in press).

Organizational Cognitive Repairs

Thus far we have emphasized interventionist approaches to modifying the environment. The "debiaser" could be a government agency, an employer, or the decision maker herself. But debiasing can also be embedded in an organization's routines and culture. Heath, Larrick, and Klayman (1998) call these debiasing organizational artifacts *cognitive repairs*. A repair could be as simple as an oft-repeated proverb that serves as a continual reminder, such as the phrase "don't confuse brains with a bull market," which cautions investors and managers to consider the base rate of success in the market before drawing conclusions about an individual investor's skill. Other examples offered by Heath et al. (1998) include institutionalizing routines in which senior managers recount stories about extreme failures (to correct for the underestimation of rare events), and presenting new ideas and plans to colleagues trained to criticize and poke holes (to overcome confirmatory biases and generate alternatives). Many successful repairs are social, taking advantage of word-of-mouth, social influence, and effective group processes that encourage and capitalize upon diverse perspectives. Although cognitive repairs may originate as a top-down intervention, many arise organically as successful practices are noticed, adopted, and propagated.

We highlight one cognitive repair that has not only improved many organizational decisions, but has also saved lives—the checklist. This tool could easily fit in many of our debiasing categories. Like linear models, checklists are a 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, by design, reduce errors due to forgetfulness and other memory distortions (e.g., over-reliance on the availability heuristic). Some checklists are so simple that they masquerade as proverbs (e.g., emergency room physicians who follow ABC—first establish airway, then breathing, then circulation, Heath et al., 1998, p.13). External 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), making them an important tool for overcoming low decision readiness. Often, checklists are reviewed socially (e.g., among a team of medical professionals), which ensures not only that best practices are followed, but also that difficult cases are discussed (Gawande, 2010).

CHOOSING A DEBIASING STRATEGY

Given that there are many available debiasing methods, what are the criteria for choosing between them? With the increased interest in policy interventions for improving a myriad of decisions, this is an important area for future research. Here we sketch six considerations that we believe are important for informing this decision: effectiveness, decision readiness, competence/benevolence, heterogeneity, decision frequency, and decision complexity.

Effectiveness

Some debiasing methods will work better than others in a given context. For example, whereas the American Cancer Society recommends that everyone over age 50 have a colonoscopy every ten years, only about half of the target population does so. Narula et al. (2013) tested two different interventions for patients between 60 and 70 years old who had

received at least one colonoscopy in the past, but for whom the recommended 10-year interval since their last screening had elapsed. Some patients were sent a letter that specified a date and time for their colonoscopy, and they had to call in to change this (an opt-out default). Others received a planning prompt—their letter reminded them that they were overdue and suggested that they call in to schedule an appointment. With the planning prompt, 85% of patients ultimately received treatment, compared to 63% in the default condition. The context undoubtedly played a role in producing this result—an upcoming colonoscopy can be distressing, and paternalistically assigning one may evoke a measure of reactance. Each context has its idiosyncrasies, and we strongly recommend that would-be choice architects consider a range of debiasing methods and run experiments to discover which is most effective. Moreover, there is also the challenge of measuring success, especially when people have heterogeneous preferences (see Ubel, 2012, for a thought-provoking discussion of possible criteria for measuring success).

Decision Readiness

In general, shortcomings in decision readiness might best be treated by modifying the environment. When people are in tempting situations or have many demands on their attention, they may lack the ability to apply many of the decision aids of the "modify the person" variety. For example, a hungry person may not pause to consider the pros and cons of loading up the plate at the dinner table. However, smaller glasses and dishes are a nudge that can help people consume less, while simultaneously circumventing the need for them to think clearly when in an unready state. Similarly, a fast-paced work environment and personal attachment to ideas may impede unbiased reflection in some organizations, and thus organizational cognitive repairs may be more successful than teaching employees about debiasing techniques for individuals.

Competence/Benevolence

The flip side of decision readiness is the competence of the prospective choice architect. Increasingly, governments and organizations 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 hold the possibility of yielding great benefits at a relatively low cost. On the other hand, modifying the environment can be problematic if policy makers mispredict individuals' preferences, or worse, have a hidden agenda. Additionally, some nudges operate below awareness, which raises the ethical question of whether it is acceptable for a policy maker to take away some individual autonomy in order to improve welfare (see Smith et al., 2013, for an illuminating discussion on this point). The more dubious the competence and benevolence of the policy maker, the more appropriate it becomes to approach debiasing by modifying the person rather than the environment.

Heterogeneity

When people vary in their preferences or biases, a given intervention could potentially leave some people worse off. Although the possibility of heterogeneity is often raised in critiques of defaults, it also has ramifications for other debiasing methods, including those that modify the person. For example, "think of con reasons" may reduce overconfidence for many, but may exacerbate underconfidence for the few individuals who are biased in that direction. To address heterogeneity, Dietvorst et al. (2014) distinguish between *outcome nudges*, which push toward a uniform outcome for all, and *process nudges*, which debias by helping individuals employ decision strategies most likely to lead to their personally preferred outcomes. Defaults are clearly outcome oriented, whereas other strategies, such as nudges that induce reflection (e.g., planned interruptions) are more process-oriented because they merely encourage people to pause and think more deliberatively about their objectives. The greater the heterogeneity, the more we should worry about "shoving" as opposed to "nudging," and the more interventions should focus on process as opposed to outcomes.

Decision Frequency

Many types of decisions are repeated, such as admitting new students to a university, investing in new businesses, or diagnosing cancer. These types of decisions provide the same

inputs (e.g., student test scores) and require the same type of response (e.g., admit or not). Linear models, checklists, and consistent policies can dramatically improve accuracy for repeated decisions. Some decisions are made infrequently by individuals but are repeated across people. Here too, models have the potential to be helpful, such as recommender systems for retirement planning that simplify choice, perhaps coupled with a dose of just-in-time financial education so that decision makers can understand the basic trade-offs they face (Fernandes et al., 2014). Finally, though, there remain many (arguably most) personal decisions big and small for which a standardized approach (if not a standardized answer) is infeasible or unavailable (e.g., choosing between a job and more education, choosing a medical treatment, deciding whether to eat out or stay in, etc.) because the specific decisions are infrequent or idiosyncratic to the individual. Modifying the person can help here. For instance, providing people with cognitive strategies to (a) identify objectives, (b) generate a broad range of alternatives, and (c) seek out disconfirming evidence, is likely to yield a high return for infrequent decisions. This can be coupled with modifying the environment, for instance by providing ample time for reflection, shaping information so that it can be understood and used appropriately, and developing routines in organizations that facilitate divergent thinking and better learning.

Decision Complexity

Many important decisions are very complex, such as choosing among dozens of available plans for health insurance or retirement savings. Even highly educated individuals sometimes have difficulty identifying the best options (Thaler & Sunstein, 2008), and some people are so overwhelmed that they do not choose (Iyengar & Lepper, 2000). To make matters worse, product complexity, as defined by number of features, is increasing in the financial services industry, which increases the likelihood of inferior choices by consumers (Célérier & Vallée, 2014). For complex decisions that are encountered infrequently (but repeated across individuals), modifying the environment via effective choice architecture is an attractive option. Moreover, if preferences are heterogeneous, we would probably want to help people navigate the terrain of options, rather

than limiting choice in some way. One promising approach for financial and health care decisions is to provide smart defaults (options pre-selected based on consumer characteristics) along with just-in-time education, and an architecture that allows for motivated consumers to explore and choose from the entire spectrum of options (Johnson, Hassin, Baker, Bajger, & Treuer, 2013).

AN EXAMPLE

Consider again the "less-now versus more-later" decision faced by retirees regarding when to begin their social security payments that we described earlier in this chapter. In the US, retirees must choose between smaller payments beginning at age 62 and larger payments beginning as late as age 70. Based on the ideas reviewed in this chapter, a variety of debiasing tools can be developed to facilitate a wise decision. As shown in Figure 1, debiasing tools can be organized from those toward the left that improve decisions by providing and shaping information, to those on the right which influence the decision making strategies that people apply. Providing completely new information is not, by itself, an example of debiasing. However, providing information counts as debiasing when it is otherwise available but tends to be neglected—the decision maker could in principle obtain the information at a relatively minimal cost. For example, the British government is considering providing life expectancy forecasts (generally available on the web) as part of a free consultation service to help retirees manage their pensions (Beinhold, 2014). Note that strategies toward the right of the spectrum presented in Figure 1 may still have an informational component (e.g., defaults might be interpreted as expert advice). The strategy on the far right of the figure involves using one's own objectives as a prompt for generating new alternatives (Hammond, Keeney, & Raiffa, 1999). For example, a new retiree who requires funds for an around-the-world vacation may discover that alternatives such as selling savings bonds or taking out a loan are financially more attractive than withdrawing money from social security early and forgoing larger payments later in life.

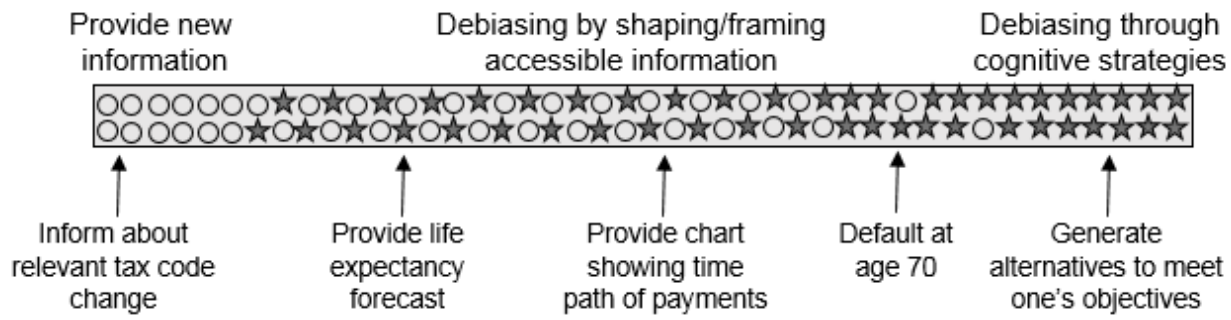

Figure 1. A continuum of debiasing strategies. By itself, new information is not debiasing, as shown on the far left. The other strategies depicted all contain elements of debiasing.

Which debiasing method is best? Although not particularly complex, choosing the start date for social security is once-in-a-lifetime decision. Moreover, decision readiness is low for the many individuals who lack basic financial knowledge or numeracy skills. These factors argue in favor of modifying the environment. On the other hand, heterogeneity in preferences suggests that a default may have the undesirable consequence of swaying some people toward an inferior choice. Other changes to the environment seem potentially helpful, such as providing a life expectancy forecast or a payment chart, assuming a competent policy maker is available to develop and implement these tools. Of course, different tools can also be combined. Prospective retirees can be provided with helpful charts, encouraged to think about the tradeoff between having extra money in their 60s versus greater resources later in life, and encouraged to consider alternative routes to meeting their financial needs.

We reiterate that potential tools should be tested experimentally to see whether they are effective. For example, a point estimate of life expectancy may be misinterpreted unless people understand the uncertainty around it. A person might react very differently to a point forecast (e.g., "our best guess is that you will live to age 81") and a range forecast (e.g., "10 out of every 100 people similar to you live to age 92 or older"). Although both forecasts might be derived from the same analysis, the latter one conveys more useful information to those who want to make sure that they have enough resources to last a lifetime.

FINAL REMARKS

Bias in judgment and decision making is a common but not insurmountable human problem. Our hope is that this review of the debiasing literature will better equip readers with a set of strategies for improving decisions (overcoming common biases) that are based on psychological principles. In many cases, however, there will be multiple reasonable options for debiasing, and therefore a need to identify the method that produces the best results. We offer six factors (and there are undoubtedly more) to consider when selecting a debiasing method. Thinking through these considerations requires an assessment of the context, and debiasing dilemmas that may emerge. For example, to whom should debiasing be entrusted: an imperfect decision maker or a fallible choice architect? We know that individuals are sometimes biased, but it is important to also recognize that policy makers can be misguided, or have interests that conflict with those of the individuals whose decisions they seek to influence. Many other such debiasing dilemmas will exist in different situations. In addition to hopefully helping people improve their decisions, and the decisions of others, we hope that this chapter stimulates future research on the important topic of debiasing. We need to increase our toolkit of potential debiasing strategies based on psychological principles, to collect evidence on what actually works in specific, context-rich environments, and finally to help people both select and use the better debiasing strategies for their particular decision problems. Regardless of whether the decisions facing an individual (or group) are professional, e.g., selecting the better employee, or personal, e.g., managing one's retirement savings and expenditures, methods for debiasing will often be needed.

REFERENCES

Acland, D., & Levy, M. (2013). Habit formation, naiveté, and projection bias in gym attendance. *Working Paper*.

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

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

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological bulletin, 110*, 486-498.

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*, 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*, 635-672.

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

Baron, J. (2012). The point of normative models in judgment and decision making. *Frontiers in psychology, 3*.

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current directions in psychological science, 16*, 351-355.

Beinhold, K. (2014, April 14). How near is the end? Britain may tell retirees. *The New York Times*. Retrieved from http://www.nytimes.com.

Ben-David, I., Graham, J. R., & Harvey, C. R. (2013). Managerial miscalibration. *The Quarterly Journal of Economics, 128*, 1547-1584.

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*.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., & Milkman, K. L. (in press). The effect of providing peer information on retirement savings decisions. *The Journal of Finance*.

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

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

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*, 1074-1078.

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

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*, 7-42.

Camilleri, A. R., & Larrick, R. P. (2014). Metric and Scale Design as Choice Architecture Tools. *Journal of Public Policy & Marketing, 33*, 108-125.

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*, 513-518. Article.

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*, 1639-1674.

Casey, B. M., McIntire, D. D., & Leveno, K. J. (2001). The continuing value of the Apgar score for the assessment of newborn infants. *New England Journal of Medicine, 344*, 467-471.

Célérier, C., & Vallée, B. (2014). What drives financial complexity? A look into the retail market for structured products. *Working Paper*.

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. *JAMA, 304*, 43-44.

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

Cheema, A., & Soman, D. (2008). The effect of partitions on controlling consumption. *Journal of Marketing Research, 45*, 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*. New York: 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*, 1-243.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International journal of forecasting, 5*, 559-583.

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

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American psychologist, 34*, 571.

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

DeSteno, D., Li, Y., Dickens, L., & Lerner, J. S. (2014). Gratitude: A tool for reducing economic impatience. *Psychological Science*. Advance online publication. Doi: 10.1177/0956797614529979

Dholakia, U. M., & Bagozzi, R. (2003). As time goes by: How goal and implementation intentions influence enactment of short-fuse behaviors. *Journal of Applied Social Psychology, 33*, 889-922.

Dietvorst, B., Milkman, K. L., & Soll, J. B. (2014). Outcome nudges and process nudges: diverse preferences call for process nudges. *Working Paper*.

Duncker, Karl (1945). On Problem Solving. Psychological Monographs 58. American Psychological Association.

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*, 398-405.

Fernandes, D., Lynch Jr, J. G., & Netemeyer, R. G. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Management Science*. Advance online publication. doi: 10.1287/mnsc.2013.1849

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

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

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

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

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

Glenn, J. C., Gordon, T. J., & Florescu, E. (2012). *2012 State of the Future*. Washington, D. C.: The Millenium Project.

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

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

Hammond, J. S., Keeney, R. L., & Raiffa, H. (1999). *Smart choices: A practical guide to making better decisions*. Boston: Harvard Business School Press.

Hammond, K. R. (1996). *Human judgement and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*: New York: Oxford University Press.

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

Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review, 112*, 494.

Heath, C., & Heath, D. (2011). *Switch*. New York: Broadway Books.

Heath, C., Larrick, R. P., & Klayman, J. (1998). Cognitive repairs: How organizational practices can compensate for individual shortcomings. *Review of Organizational Behavior*, 1-37.

Herzog, S. M., & Hertwig, R. (2014). Think twice and then: Combining or choosing in dialectical bootstrapping? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 218-232.

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science, 290*, 2261-2262.

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

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes, 67*, 247-257.

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*, 995-1006.

Jain, K., Mukherjee, K., Bearden, J. N., & Gaba, A. (2013). Unpacking the future: A nudge toward wider subjective confidence intervals. *Management Science, 59*, 1970-1987.

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

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

Johnson, E. J., Hassin, R., Baker, T., Bajger, A. T., & Treuer, G. (2013). Can consumers make affordable care affordable? The value of choice architecture. *PLOS ONE, 8*.

Johnson, E. J., Shu, S.B. , Dellaert, B. G. C., Fox, C.R., Goldstein, D.G., Haubl, G., Larrick, R.P., Payne, J. W., Peters, E., Schkade, D., Wansink, B., & Weber, E. U. (2012). Beyond nudges: Tools of choice architecture. *Marketing Letters, 23*, 487-504.

Peters, E., Schkade, D., Wansink, B., and Weber, E. U. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, 23, 487-504.

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

Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267-293).

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-291.

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

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

Klayman, J., & Brown, K. (1993). Debias the environment instead of the judge: An alternative approach to reducing error in diagnostic (and other) judgment. *Cognition, 49*, 97-122.

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

Larrick, R. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making.* Malden, MA: Blackwell.

Larrick, R. P. (2009). Broaden the decision frame to make effective decisions. In E. A. Locke (Ed.), *Handbook of Principles of Organizational Behavior (2nd ed.)*. Chichester, UK: Wiley and Sons.

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*. New York, NY: Psychology Press.

Larrick, R. P., Morgan, J. N., & Nisbett, R. E. (1990). Teaching the use of cost-benefit reasoning in everyday life. *Psychological Science, 1*, 362-370.

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

Larrick, R. P., & Soll, J. B. (2012). *Combining the intuitive and analytic mind*. Paper presented at the Behavioral Decision Research in Management, Boulder, CO.

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

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125*, 255-275.

Lerner, J. S., Li, Y., & Weber, E. U. (2013). The financial costs of sadness. *Psychological Science, 24*, 72-79.

Lewis, M. (2003). *Moneyball: The art of winning an unfair game*. New York: W.W. Norton & Company.

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

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*, 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*, 51-68.

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

McKenzie, C. R. M., Liercsch, M. J., & Finkelstein, S. R. (2006). Recommendations in policy defaults. *Psychological Science, 17*, 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*, 163-176.

Milkman, K. L., Akinola, M., & Chugh, D. (2012). Temporal distance and discrimination: An audit study in academia. *Psychological Science, 23*, 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*, 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*, 92-93.

Milkman, K. L., Minson, J. A., & Volpp, K. G. (2013). Holding the hunger games hostage at the gym: An evaluation of temptation bundling. *Management Science*. Advance online publication.

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*, 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*, 17-35.

Milne, S., Orbell, S., & Sheeran, P. (2002). Combining motivational and volitional interventions to promote exercise participation: Protection motivation theory and implementation intentions. *British Journal of Health Psychology, 7*, 163-184.

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*, 25-38.

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

Narula, T., Ramprasad, C., Ruggs, E. N., & Hebl, M. R. (2013). Increasing colonoscopies? A psychological perspective on opting in versus opting out. *Health Psychology*. Advance online publication.

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

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

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

Prelec, D., & Loewenstein, G. (1991). Decision making over time and under uncertainty: A common approach. *Management Science, 37*, 770-786.

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*, 189-205.

Rothman, A. J., & Salovey, P. (1997). Shaping perceptions to motivate healthy behavior: the role of message framing. *Psychological bulletin, 121*, 3-19.

Rothman, A. J., & Updegraff, J. A. (2010). Specifying when and how gain-and loss-framed messages motivate healthy behavior: An integrated approach. In G. Keren (Ed.), *Perspectives on framing*. London: Psychology Press.

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

Russo, J. E. (1977). The Value of Unit Price Information. *Journal of Marketing Research, 14*, 193-201.

Russo, J. E., & Schoemaker, P. J. H. (2013). Overconfidence. In M. Augier & D. J. Teece (Eds.), *The Palgrave Encyclopedia of Strategic Management*.

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

Smith, N. C., Goldstein, D. G., & Johnson, E. J. (2013). Choice without awareness: ethical and policy implications of defaults. *Journal of Public Policy & Marketing, 32*, 159-172.

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

Soll, J. B., Keeney, R. L., & Larrick, R. P. (2013). Consumer misunderstanding of credit card use, payments, and debt: causes and solutions. *Journal of Public Policy & Marketing, 32*, 66-81.

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

Soman, D., & Cheema, A. (2011). Earmarking and partitioning: increasing saving by low-income households. *Journal of Marketing Research, 48*, S14-S22.

Soman, D., Xu, J., & Cheema, A. (2010). Decision points: A theory emerges. *Rotman Magazine* (Winter): 64-68.

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

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

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. London: Little, Brown.

Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy, 112*, 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*, 1275-1289.

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

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

Ubel, P. A. (2013). Beyond knowledge: Figuring out how to help people make "good" decisions. In E. Shafir (Ed.), *The Behavioral Foundations of Public Policy*. Princeton, NJ: Princeton University Press.

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

Wansink, B., & Cheney, M. M. (2005). Super bowls: Serving bowl size and food consumption. *JAMA, 293*, 1723-1728.

West, R. F., Meserve, R. J., & Stanovich, K. E. (2012). Cognitive Sophistication Does Not Attenuate the Bias Blind Spot. *Journal of Personality and Social Psychology, 103*, 506-519.