

What Does It Mean to Be a Rational Decision Maker?

Barry Schwartz

[Paper submitted for publication 08/01/13]

## What Does It Mean to Be a Rational Decision Maker?

Barry Schwartz

### Abstract

Research on the psychology of decision making has historically relied on the principles of rational choice theory to provide a normative standard. For the most part, empirical research has documented deviations from this normative standard, with debate often centered on just how costly to individuals these deviations are. This paper challenges several normative features of the rational choice model. It suggests that “maximizing” (of utility, expected value, satisfaction) is often not the appropriate normative goal. It suggests that the value of decision outcomes cannot be assessed independent of the decision contexts that give rise to them (i.e., that the value of outcomes is not “path independent.”). And it suggests that the relation between the magnitude of an outcome (or a psychological characteristic) and its value is often non-monotonic. I argue that the honorific “rational” should be based on the substantive and not the formal properties of decisions—that an adequate theory of rational decision making should consider the way in which decisions enable people to live good, meaningful, and satisfying lives. And understood in this way, the hallmark of rationality is wise judgment.

Key words: Rational choice theory, maximizing and robust satisficing, “leaky” rationality, inverted U-shaped functions

## What Does It Mean to Be a Rational Decision Maker?

Barry Schwartz

Swarthmore College

“Respondents offer a reasonable answer to a question they have not been asked... respondents who substitute one attribute for another are not confused about the question they are trying to answer—they simply fail to notice that they are answering a different one” (Kahneman, 2003, p.709).

In the thirty years or so since it began, the field of behavioral decision making, or judgment and decision making (JDM), has developed an ever-growing catalogue of the mistakes human beings are susceptible to when they use a variety of heuristics and biases to evaluate information, make decisions, and then evaluate the results of those decisions. In assessing probability, people seem to interpret “how likely is this event to happen?” as “how typical is this of the class of events of which it is a member?” People treat the vividness of an event in memory as an indication of how frequently the event occurred in the past. People make risk averse choices when choosing among possible gains and risk seeking choices when choosing among possible losses. This is not, in itself, a problem, but it becomes a problem when variations in the language of description can induce people to treat the identical choice situation as one involving gains or as one involving losses. People organize inflows and outputs of money into a variety of mental accounts,

which helps explain why they are willing to treat themselves to a luxury when they have a windfall, but otherwise not. This also helps explain why people will make deposits into savings accounts that pay 3% interest while at the same time making minimal payments to reduce credit card debt at 18% interest. Peoples' assessments of the value of a good at a given price are dependent on surrounding other goods that provide "anchors" (eg., a \$600 suit may be a "steal" on a rack of \$1000 suits, but an extravagance on a rack of \$300 suits). Phenomena like these have grown out of the research program on heuristics and biases launched by Daniel Kahneman and Amos Tversky (e.g., Gilovich, Griffin, & Kahneman, 2002; Kahneman, 2003, 2011; Kahneman & Tversky, 1984, 2000). And they have led to a kind of "two-process" theory of judgment and decision making. One process, which is rapid, automatic, and inaccessible to consciousness, delivers results to consciousness that are produced by these heuristics. Afterwards, the second, slower process, which is conscious, effortful, and rule-governed, may go to work using logic, probability theory, and other formal systems. A decision maker need not accept the results of the automatic process as competent or definitive, but this process delivers answers upon which consciousness acts. The results of the operation of the heuristics and biases of the automatic process do not always lead to mistaken judgments and bad decisions. Indeed much of the time, they serve us well (see, e.g., Gigerenzer, 2007). Nonetheless, thirty years of research documents that sometimes, they can lead to serious errors.

In all the research on how heuristics and biases can lead people into bad decisions, the normative standard for comparison has rarely been called into question. However, in this paper, I will argue that many decisions we face cannot be handled by the

formal systems that are taken for granted as normatively appropriate. Kahneman has come to see much of the work he did with Amos Tversky as revealing a process of “attribute substitution” (Kahneman 2003, 2011; Kahneman & Frederick, 2002) in which people asked a hard question (eg., “how likely is X?”) answer an easier one (“how typical is X?”) without realizing that the question they have answered is different from the one they were asked. The theme of this paper is that researchers in judgment and decision making have done the same thing. The world wants to know what it means to make a decision that is *substantively* rational (i.e., will enable people to achieve their goals and lead successful and satisfying lives) and what the impediments to substantive rationality are. Decision making research has answered an easier question, namely, what it means to make a decision that is *formally* rational and what the impediments to *that* are. In this paper, I will suggest that substantive and formal rationality are different, and that at least some tenets of formal rationality lead to decisions that are sub-optimal with respect to substantive rationality. The arguments below are meant to be suggestive, not definitive; illustrative, not exhaustive. My aim is to encourage JDM researchers to start trying to answer the question that the world is actually asking.

### The Normative Theory of Rational Choice

Let me set the stage for the argument by sketching some key principles of the theory of rational choice. The presumed goal of decisions is to maximize utility, or preference. The assumption is that people bring well-articulated preferences with them to the decision space (i.e., that preferences are “exogenous”). People then array the options before them, analyzing them into relevant attributes, and assign importance weights to

each attribute. They then assign values on each attribute to each alternative as well as probabilities that if they choose a given alternative, their goal with respect to the target attribute will be realized. Then, rational choosers just do the math, and choose the alternative with the highest value. Sometimes, there will be a dominant alternative so that this elaborate apparatus will seem excessively laborious and cumbersome. But much of the time, tradeoffs will have to be made, and this apparatus may prove essential.

This description is admittedly highly schematic, but it enables me to highlight a few key points. First, the structure of rational choice theory is entirely formal; one could substitute variables for actual alternatives and attributes and have a recipe that applies to all decisions. Second, deviations from this normative model will also be formal. That is, inconsistency, intransitivity, preference reversals, lack of descriptive invariance, and the other phenomena studied by JDM researchers are all identified as “errors” by their failure to match the formal, normative model. Finally, rationality is judged for the circumscribed decision itself and not by assessing how the particular decision fits into the decider’s life as a whole.

Everyone knows that the normative theory is an idealized abstraction that does not approximate how decisions are actually made, or even how they should be made. Going through the process of decision analysis may be more costly in time and cognitive resources than the decision is worth. And an outcome that is utility maximizing in an individual decision may be destructive when cumulated, so that individual decisions must be considered in terms of the long-term consequences they may have. This acknowledgement has led some researchers, in the spirit of Herbert Simon (1955, 1956, 1957) to modify the rational choice norm and speak of “bounded rationality,” which

highlights the cognitive (and emotional) limitations of human beings. The notion of bounded rationality leaves the normative status of the model of rational choice intact, and simply describes the ways in which finite organisms actually make decisions with processes that fall short of the normative standard. Thus, I think the normative standard exerts a powerful influence on the research that is actually done, on what investigators find interesting and noteworthy, and on the prescriptions that are offered to improve decision making. Perhaps most significant, the normative standard makes certain important questions about rationality essentially invisible to researchers and policy makers alike. My focus is not on whether people adhere to the normative theory; they do not. Rather, my focus is on whether the normative standard *should* be normative. It is to this issue that I now turn.

The following discussion has three distinct arguments. The first is that there are circumstances, almost certainly quite common, in which “maximizing” preference or utility is not the normatively correct thing to do. Rather, the appropriate norm is what has been called “robust satisficing” (Ben-Haim, 2006; Schwartz, Ben-Haim, & Dacso, 2011). The second is that because the context of choice often affects the chooser’s experience with the option that is chosen, and because “utility” or “preference” is subjective, it is commonly the case that decisions that violate the formal norms of rational choice will be vindicated by the chooser’s experience. On such occasions, the only stance from which to judge the rationality of such decisions is one that evaluates the substantive effects of the decisions on the chooser’s life. The third is that there are many personal attributes and environmental variables with effects on behavior that are non-monotonic. More of variable or attribute X improves behavior Y until a point is reached

at which the curve relating X to Y changes direction. Where and why the curve bends is context specific, so that it is the substantive, and not the formal properties of a situation that determine when there is “too much of a good thing.” In the case of each of these arguments, we will see, and judgment of the rationality of the decision process will require an assessment of the substantive effects of the decision on the chooser, rather than an assessment of the formal or logical properties of the decision-making process.

The foregoing preview is somewhat abstract, so I will make it concrete with a brief review of research that I and others have done on choice overload and maximizing. This body of research is the source of each of the three arguments I will be making.

### Choice Overload and Maximizing

Though the logic of rational choice theory would seem to imply that more options are always better than fewer (since people can elect to ignore options if they are not interested in exploring them), Iyengar and Lepper (2000) demonstrated that when a choice set is large enough, the likelihood that people will choose at all goes down (see Botti & Iyengar, 2004, 2006; Iyengar, 2010; Iyengar, Jiang, & Huberman, 2004; Schwartz, 2004). Thus, the relation between choice set size and the likelihood of choice is non-monotonic (but see Scheibehenne, Greifeneder, & Todd, 2010, for a discussion of the limits of this effect). Moreover, large choice sets seem to lead to worse decisions (eg., Hanoch, Wood, Barnes, Liu, & Rice, 2011), and to dissatisfaction with even good decisions (Iyengar & Lepper, 2000; Schwartz, 2004). Though there is some uncertainty about how robust this “choice overload” phenomenon is, and about the conditions under which it occurs, there is no doubt that it is a real phenomenon.

This capsule description of choice overload illustrates two of the three arguments I will make in this paper. First, the relation between choice set size and utility is non-monotonic; that is, choice is good, but there can be too much of a good thing. And second, the conditions of choice (ie., a large vs. a small choice set) affect satisfaction with what is chosen.

The third of my arguments derives from research I did with colleagues that shows that the “choice problem” is especially acute for people who aim to maximize in their decisions (Iyengar, Wells & Schwartz, 2006; Schwartz, 2004; Schwartz, Ward, Lyubomirsky, Monterosso, White, & Lehman, 2002). My colleagues and I developed a scale intended to measure individual differences in tendencies to maximize (get the “best”) as opposed to satisfice (get “good enough”). We found that there are such differences and that people who score high on our measure of maximizing tendencies have the most difficulty choosing, want the most alternatives, are most willing to devote resources (in time and/or money) to expanding their choice sets, and are most likely to compare their choices to those of others. In addition, we found that although maximizers often do better than satisficers on objective measures of decision quality, they feel worse about how they do (see also Dar Nimrod, Rawn, Lehman, & Schwartz, 2009; Polman, 2010; and Sparks, Ehrlinger, & Eilbach, 2012 for further evidence, but see Diab, Gillespie, & Highhouse, 2008, for a critical evaluation of the scale). This effect of maximizing on decision satisfaction suggested to us that a maximizing strategy, though normatively appropriate, was psychologically counterproductive. If the aim of decisions is to increase utility, which is a subjective entity, then doing better, by maximizing, fails if it makes the decision maker feel worse.

My initial interpretation of this literature on choice overload and maximizing was that they were further examples of “bounded rationality.” That is, I thought the normative presumptions that more choice is better than less and that utility maximization should be the goal of decisions was not challenged by this work. What the work showed was only that people are unable to behave in choice situations as fully rational actors. What I now think, and what the rest of this paper will try to argue, is that choice overload and maximizing results are not instances of the limits of human rationality; rather, they are instances of the limits of rational choice theory as a normative model of what rationality means.

The remainder of the paper will argue that:

1. There are circumstances in which satisficing, not maximizing is the right normative strategy.
2. The context of choice routinely “leaks” into the experience of the results of that choice, and this leakage creates problems for normative theories of rational choice that focus on the formal properties of decisions.
3. As with choice set size, there are many areas in life where more of something good can be worse than less, requiring judgment about how much of that good thing to pursue. This leads to the view that rationality is best understood as good judgment rather than adherence to a set of formal or logical rules.

### Maximizing, Satisficing and “Radical Uncertainty”

Suppose you’ve been fortunate enough to be admitted to a half-dozen colleges. Now, you sit down to decide which one to attend. How should you go about this process?

It is generally agreed that the best approach is to do a multi-attribute utility analysis (Keeney & Raiffa, 1993.) First, put together a big spreadsheet. Then, list all the things that matter to you about college (e.g., size, location, reputation, quality of its program in field biology, social life, music department, housing, etc.) Then, attach a weight to each attribute, to reflect its importance to you. If you are devoted to field biology, it may get a weight of 1.0, while other dimensions get fractions of that weight. Next, evaluate each school on each dimension; give it a score, say from 1–10. Finally, multiply scores by weights, and do some addition. Choose the school with the highest score.

This process can obviously be taxing and time consuming, but the situation is even more complex. When you assign scores for each school on each dimension, you're making guesses or predictions. Your assessment of the music department, the field biology program, and the social life may be wrong. So to acknowledge uncertainty, you will need to assign probabilities to the values in each cell of the spreadsheet. And the situation is more complex still. You may be wrong about how important field biology, social life, and location are to you. You're only seventeen, after all, and people change. So the weights you attach to dimensions also need probabilities. There is an additional complexity. Even if your estimates of importance and quality are correct, you don't know how it will actually feel to experience being a student at a school that has the qualities of the one you choose. You are making a prediction about a future subjective state, and as Daniel Gilbert, Timothy Wilson, and their various collaborators have amply documented, (e.g., Gilbert, 2006; Wilson & Gilbert, 2005), such predictions are notoriously inaccurate. And there is one final matter. There are some influences on your satisfaction with college that just can't be predicted. Will you get along with your

roommate? Will the best professor in the biology department leave? Will you form a romantic attachment? These kinds of factors can play a major role in determining your college experience, and they are inherently uncertain. You can't even pretend to attach probabilities to them, or even to identify all of them. Making this decision is tough. You could easily be wrong. Nonetheless, you do the best you can, and that seems to be multi-attribute utility calculation. It's your best strategy.

Or is it? Imagine a different goal. Given the multiple sources of uncertainty that are a part of the process, suppose your goal is to choose the school that is likely to be satisfactory, even if your estimates of its quality and probability on various dimensions are wrong. Instead of maximizing utility if everything goes well, you are trying to maximize confidence in an acceptable outcome, even if you suffer the slings and arrows of outrageous fortune. Ben-Haim (2006) calls such a goal "robust satisficing" ( see also Ben-Haim, 2011; Schwartz, Dacso, & Ben Haim, 2011). You are still trying to maximize something, but what you're trying to maximize is your confidence in a good enough outcome even if things go poorly. There is no particular reason to assume that the school that is best in your utility calculation is also the school that is most robust to error in the data underlying that calculation.

*Radical Uncertainty.* What this scenario, and countless others (e.g., buying a car, choosing a place to go on vacation; choosing a job; choosing a treatment plan for a serious medical condition; choosing investments for your retirement), have in common is that you are faced with a decision that has multiple dimensions, with outcomes that are uncertain and influenced by factors that are difficult to evaluate or even identify. And they are not merely uncertain in a probabilistic sense. In many cases, you cannot even

attach probabilities in a meaningful way. Your uncertainty is more radical than the uncertainty you face when rolling dice. Knight (1921) distinguished between probabilistic risk and non-probabilistic “true uncertainty,” as he called it. Ellsberg (1961) famously pointed out this distinction when he contrasted an urn with 50 red and 50 black balls with an urn that has 100 balls, some of which are red and some black. If their task is to pick a red ball, people typically prefer the first urn to the second, preferring (probabilistic) uncertainty to what Ellsberg termed “ambiguity.” The thrust of my advocacy of robust satisficing as a decision criterion is this:

1. Most of the decisions people face in life involve Knightian uncertainty or ambiguity at least as much as they involve probabilistic uncertainty. This is especially true when a key feature of a decision is the person’s estimation of how it will feel to have one outcome rather than another. For example, experiencing a side effect (e.g., impotence) of prostate cancer surgery is one thing; estimating the subjective consequence of this side effect, before the fact, is quite another.

2. In conditions of radical uncertainty, utility maximization as a strategy is unreliable. Indeed, it may even be self-deceptive, in that it involves assigning probabilities to outcomes in a context in which probabilities cannot be specified.

3. There is a quite reasonable alternative to utility maximization. It is maximizing the robustness to uncertainty of a satisfactory outcome, or robust satisficing. Robust satisficing is particularly apt when probabilities are not known, or are known imprecisely. The maximizer of utility seeks the answer to a single question: which option provides the highest subjective expected utility. The robust satisficer seeks to answer *two* questions: first, what will be a “good enough” or satisfactory outcome; and second, of the options

that will produce a good enough outcome, which one will do so under the widest range of possible future states of the world.

4. This alternative has been formalized as “info-gap decision theory” (Ben-Haim, 2006). Though I will not discuss the theory here, it has been used effectively as a decision-making framework in an extremely wide variety of domains, though none of them, to date, are psychological (see Ben-Haim, 2006, 2011 for the formal theory and Schwartz et al., 2011, for more discussion and references).

Info-gap decision theory is designed to handle situations of profound uncertainty. Since we do not know how wrong our data and models are, we evaluate a proposed decision by asking: what is the greatest horizon of uncertainty at which the decision will still yield acceptable results? How wrong can we be, in our understanding of the relevant processes and requirements, and the outcome of the decision still be acceptable? The answer to this question is the *robustness function*. The robustness function generates a preference ordering on the available options: a more robust option is preferred to a less robust option. Satisficing means doing well enough, or obtaining an adequate outcome. A *satisficing decision strategy* seeks a decision whose outcome is good enough, though perhaps sub-optimal. A *robust-satisficing decision strategy* maximizes the robustness to uncertainty and satisfices the outcome.

What are the conditions to which robust satisficing applies? Answering this question requires a brief excursion into the foundations of probability theory. Baron (2008, and see Brown, 1993) nicely summarizes three different approaches to understanding what probability statements mean. The first, we might call “logical.” When the events that comprise a sample space are fully known, and their distributions

can be specified, a probability statement is simply a matter of logic: in the sample space of outcomes of rolls of two dice, there are 36 equi-probable outcomes, of which six sum to “7.” Thus one-sixth of possible rolls ( $.17$ ) will yield the outcome of interest. This is not an empirical matter. It is part of what it means to be throwing “fair” dice.

The second, we might call “empirical.” If you follow a sample of 10,000 men between the ages of, say 40 and 75, and 300 of them develop prostate cancer, you might infer that the chances of any particular man developing prostate cancer are  $300/10,000$ , or  $.03$ . You use the frequency of the event of interest in the past to infer the probability of the event with respect to any particular case in the future.

The final approach to probability we might call “personal” (see Savage, 1954). You are asked, in April, “will Spain win the World Cup this year?” “I think they will,” you say. “How sure are you?” “I give them a 25% chance,” you say. Because each World Cup competition is a unique event in ways that matter to prediction, you can’t really rely on frequencies in the past to infer probabilities in the future. The number you supply is merely an expression of your confidence. As Baron (2008) points out, some have argued that it makes no sense to attach probabilities to unique events. But, of course, each throw of the dice is a “unique event,” and each middle-aged man is a “unique event,” so distinctions among these three approaches to understanding probability statements are not so easy to make sharply. This is especially true when it comes to distinguishing frequency and personal approaches to probability.

So what, then, does it mean to call an event “radically uncertain” in a way that distinguishes throwing dice from choosing a college? What makes attaching probabilities to varying degrees of satisfaction with a college’s biology program different from

predicting the weather? It might be that if you pushed a high school senior, she would attach a number to how likely she was to love biology at Yale. But would the number mean anything? And if not, is there information available so that if she collected it assiduously, the number she attached *would* mean something? Even if the answer to this latter question is “yes,” if the meaning of the number is not entirely resolved by the added information, then there is radical uncertainty.

It seems unlikely that there will ever be models of satisfaction with college that approximate the predictive power of, say, meteorology, but that is an empirical question. There is no doubt that people can know more or less about a domain in question, so that estimates of probability from frequency can be more or less well justified. In other words, in real-life decisions, we may never be confronted with the kind of uncertainty we face with Ellsberg’s urn, where any number of red balls, from 0 to 100, is possible. But before we attach probabilities to outcomes, we need to assess which of Ellsberg’s two urns the decision we face more closely resembles.

From the point of view of normative theories of rational decision making, the key idea here is that one cannot do a conventional utility analysis without attaching probabilities to various outcomes. Inventing probabilities in the face of serious information gaps, because you have learned that that is the normatively correct way to make decisions, can lead you astray. Info-gap robust satisficing provides a rational alternative to the norms of rational choice theory.

What is challenging about this account is that for many, there is no way to think about uncertainty aside from using probability. So to say that the field biology teacher *might* leave Yale is just to say that there is some probability of departure, even if we do

not and cannot, know what that probability is. But, in fact, info-gap models are non-probabilistic (for technical details see Ben-Haim, 2006). They entail no assumptions about or choice of a probability distribution. They do not even entail the presumption that a probability distribution exists. A robust satisficing decision is one whose outcome is acceptable for the widest range of possible errors in the best estimate of relevant future states of the world. No probability is presumed or employed.

*Robust Satisficing: Normative or Prescriptive?* When Simon (1955, 1956, 1957) first introduced the term “satisficing,” he was making a prescriptive argument. The alternative to satisficing—utility maximizing—was not feasible, given the limits of human cognition and the complexity of the environment. An “ideal” human, with unlimited capacity, should maximize, but for an actual human, it would usually be a foolhardy undertaking. It is important to emphasize here that whereas Simon’s formulations were focused on the processing limitations of organisms, my discussion is focused on epistemic uncertainties inherent in the environment in which decisions get made. No amount of information-processing capacity will overcome a decision space in which probabilities—whether of outcomes, or of people’s subjective responses to outcomes—cannot be specified. Under these conditions, it is my view that maximizing robustness to uncertainty of a good enough outcome is the appropriate norm. Maximizing expected utility is not, not least because one can’t really compute expected utilities.

Thus, when people are faced with decisions under radical uncertainty, the normative theory of rational choice not only fails to give an accurate account of how people *do* choose; it gives an inaccurate account of how people *should* choose.

### “Leaky” Rationality

“In [some] cases, the framing of decisions affects not only decision but experience as well... In such cases, the evaluation of outcomes in the context of decisions not only anticipates experience but also molds it.”

—Kahneman & Tversky, 1984, p. 48

Imagine this. You’ve just been awarded tenure, and friends have taken you to an elegant, very expensive restaurant to celebrate. The menu has more than a dozen entrees, and as you read it, each option sounds more exquisite than the last. You know you can’t go wrong, no matter what you choose, but you also know that you won’t have another chance until you get promoted to full professor, if ever.

So you study the menu. On and on it goes; each dish enticing, but no clear winner. Eventually you will choose, and a sumptuous, inviting dish will be placed before you. Will you put your anguished decision making behind you and enjoy what you’ve chosen thoroughly and completely, or will your decision conflict linger, diminishing the satisfaction you ought to be getting from your delicious meal? In other words, do the processes by which people make decisions “leak” into their experience of those decisions? In a paper with Daniel Keys (Keys & Schwartz, 2007), I argued that the answer to this question is often yes, and that an appreciation of this fact has profound implications when we try to evaluate the decisions people make against the demanding standards of normative rationality. And whereas the field of judgment and decision making has always “known” that decisions leak into experience (witness the Kahneman and Tversky quote above), inadequate attention has been paid to the empirical dimensions and the theoretical implications of such leakage.

Consider the following, classic, hypothetical situations:

1. Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences are as follows

Option A: If Program A is adopted, 200 people will be saved.

Option B: If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved.

2. Option C: If Program C is adopted, 400 people will die.

Option D: If Program D is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die.

3. You travel into the city by train to see a concert, and when you get to the box office and reach into your pocket for money with which to buy a \$20 ticket, you discover that you lost a \$20 bill. Assuming that you have enough money to cover the costs, do you

A. Buy a ticket anyway

B. Go home and not see the concert

4. You travel into the city by train to see a concert, and when you get to the box office and reach into your pocket for the \$20 ticket, you discover that you lost it somewhere en

route to the concert. When you explain your predicament at the box office, the attendant is very sympathetic, but explains that since there is no way of proving that you had a ticket, and seats are unreserved, if you want to see the concert, you will have to buy another ticket. Assuming that you have enough money to cover the costs, do you

- C. Buy another ticket
- D. Go home and not see the concert

According to rational choice theory, there is a right way and a wrong way to answer these questions. Examples 2 and 4, are identical to Examples 1 and 3 in all relevant respects, though the situations have been “reframed” so that some superficial details differ. Options (C) and (D) are essentially the same as options (A) and (B) in each pair of questions, so anyone who gives answers (A) and (D), or answers (B) and (C), to consecutive questions is acting in violation of the normative principles of rationality. But determining which changes to a situation are relevant to a decision and which are merely superficial “reframing” is not a simple task, and in scenarios like these, an apparently trivial change can have real consequences for the life of the decision maker.

The Asian disease problem, presented here in Examples 1 and 2, was devised by Tversky and Kahneman (1981; see also Kahneman & Tversky, 1984) as an example of “invariance.” The principle of invariance states that changes in the descriptions of outcomes should not alter one’s preference order. Invariance is such an obvious principle that many accounts that attempt to formalize the rules of rational decision making use it implicitly, without stating it explicitly. It is obvious that the two Asian disease problems are identical when they are viewed side by side. However, when answered separately,

many people select options A (72%) and D (78%). This pattern of results violates the principle of invariance, but people make these choices because they think in the way that has been elegantly described by prospect theory (Kahneman & Tversky, 1979; Kahneman, 2003). Because people have diminishing marginal sensitivity to both gains and losses, they tend to be risk seeking when dealing with losses (people who die) but risk averse when dealing with gains (people saved).

The third and fourth hypothetical situations come from the literature on mental accounting (e.g. Kahneman & Tversky, 1984; Thaler, 1980; 1999). You are in the same financial situation in both cases. Either you could be out \$40 and see the show, or you could be out \$20 and not see the show. However, people do not take this kind of global perspective on their money. Instead, they group their finances into many narrower accounts (and see Sunstein, Kahneman, Schkade, & Ritov, 2002, for evidence that narrow accounting extends far beyond matters of personal finance), and they make decisions based on evaluations of the relevant mental accounts. People who have lost their ticket post both the \$20 cost of the first ticket and the \$20 that would be necessary to buy a new ticket to the “concert” account, which makes the total cost of attending the show seem to be \$40. Since \$40 is a lot to pay, many people who receive this version of the scenario (42% in the study by Kahneman & Tversky, 1984) do not buy another ticket. People who have lost a \$20 bill account for that loss in a way that is unrelated to the concert, so most of them (90% in the study by Kahneman & Tversky, 1984) choose to pay for a ticket. Rational choice theorists tend to take the global perspective on people’s finances and view the two situations as identical in all relevant respects. Money is fungible—one dollar is the same as any other—so wasting \$20 on a lost ticket is the same

as wasting \$20 on a lost \$20 bill. The inconsistency between the two situations caused by mental accounting, then, is another violation of the principles of rationality.

The irrationality of this pattern of responses seems obvious, but Keys and I argued that there are circumstances in which the rational course of action seems to be contrary to these obvious normative principles. The general structure of our argument was this:

1. In many cases in which people decide among alternatives, it is the *subjective* rather than the *objective* consequences of the decision that should be the standard for assessing the rationality of the decision.

2. The circumstances under which the decision is made (e.g. the way in which the alternatives are presented), affects the way in which the consequences of the decision are experienced.

3. What this means is that various “failures” of rationality that are well-studied and well-understood as determinants of a decision continue to exert their influence after the decision is made. In other words, the decision making process “leaks” into the subsequent experience of the results of the decision.

4. Whenever points 1-3 above hold, a normative assessment of the rationality of a decision will have to be substantive rather than formal. The assessment will have to focus on how the consequences of the decision are experienced, and on how the decision-making process fits into the decision maker’s life as a whole.

*“Leakage” into Experience.* Consider this example: Sally’s bosses tell her that they appreciate the work she has done during the past year, and give her a \$1000 bonus. On the same day, she receives a \$100 tax refund from the IRS. Erica’s bosses also tell her that they appreciate the work that she has done during the past year, but she is only given

a \$100 bonus. On the same day, she receives a \$1000 refund from the IRS. The situation resembles Examples 3 and 4 above, where the amount of money involved in two situations was identical but mental accounting led people to treat them differently. For Sally and Erica, though, the difference between the two scenarios seems more meaningful. A larger bonus may be more of a sign that her boss appreciates her work, and more cause for celebration. Thus, it does not seem inappropriate for Sally and Erica to react differently to the \$1,100 that they gained, or even to spend the money differently, with Sally perhaps spending more of it on celebration. In this case, there does not seem to be much of a temptation to assert that different behaviors by Sally and Erica would be a violation of formal principles of rationality.

Although it is easy to identify differences between this example and the concert ticket example that could allow the mental accounting “framing effect” to be rational in one case but not in the other, it is much more difficult to give a systematic account of these differences. To state the point more generally, even if principles like invariance seem completely uncontroversial, there is often a problem in determining when they apply. There are many situations that closely resemble a case where a normative principle is thought to apply where it nonetheless seems clear that the principle should not be followed.

One seemingly straightforward account of the application of formal principles like invariance is that they only apply when the outcomes they count as the same are in fact *experienced* as the same, and that any time outcomes are *experienced* as different, they must be treated as different, so that the formal rules simply don’t apply. Kahneman (2003) embraces this idea when he offers a way for researchers to sidestep the problem of

determining what counts as a violation of formal normative rules: let decision makers set their own normative standards. Describing the history of his work on framing effects, Kahneman (2003) reports that “the question of how to determine whether two decision problems are the same or different does not have a general answer. To avoid this issue, Tversky and I restricted the definition of framing effects to discrepancies between choice problems that decision makers, upon reflection, consider effectively identical” (p. 702). Kahneman’s approach shelters researchers from the theoretical difficulties of applying formal principles like invariance to particular, concrete situations. And he implies that the framing effects that he and Tversky studied meet the standard of reflection that he gives for identifying equivalent scenarios.

However, there has been little published research documenting participants’ reflective judgments on framing problems. Frisch (1993), in one of the few studies that systematically investigated this question, found that participants who reflect on these types of problems often insist that there are real differences between the two versions of a scenario. Frisch presented participants with fifteen pairs of scenarios like those described above (the members of each pair were not presented side by side). After participants had chosen their responses to all 30 questions (two versions of each of the fifteen scenarios), participants were shown both versions of each scenario side by side and asked a yes or no question: “Do you think the difference in these two situations warrants treating them differently?”

In 14 of the 15 problems, a majority of the participants who showed the expected framing effect answered that yes, the situations do warrant different treatment. Only on the Asian disease problem did a majority of participants favor treating the two scenarios

the same, but even on this seemingly straightforward violation of invariance only 21 out of 40 participants (53%) took this position.

Frisch (1993) argued that her results should be taken as evidence that people who are subject to framing effects are not always acting irrationally. She presented her empirical studies as a way of deciding between two accounts of framing effects. The first account is that framing effects are like perceptual illusions that trick decision makers, with framing affecting decisions but not subsequent experiences. The alternative account, which she attempted to defend, is that “framing has an effect on decisions because it has an effect on experience” (p. 402). She used participants’ reflective judgments about the scenarios as an indirect means to investigate whether framing effects influence people’s subsequent experiences. When participants were asked to justify their reflective judgment, she found, for every type of scenario, that many of their justifications were assertions that there *is* a subjective difference between the two versions of a given scenario.

There is additional, related evidence. For instance, yogurt labeled “95% fat free” tastes less rich than yogurt labeled “only 5% fat,” (Sanford et al., 2002). Beef that is “75% lean” tastes better than beef that is “25% fat” (Levin & Gaeth, 1988). There are also counterintuitive cases where seemingly dominant circumstances lead to an outcome that is worse in some ways. People have been found to perform better after drinking an energy drink that cost \$1.89 than after getting the same energy drink for \$0.89 (Shiv, Carmon, & Ariely, 2005), and to have less negative memories of an unpleasant experience if it continues longer at a reduced level of unpleasantness (e.g., Kahneman, 2000).

Frisch's findings make a simple but profound point. Changes to the description of a set of options should not be ignored if they affect people's experiences, and almost any change to a situation that a decision maker faces, even a change that seems superficial and irrelevant, may end up having some effect on the decision maker's experiences. Even in scenarios that are designed to manipulate the decision maker's choice without changing the outcomes that result from the different options, the alterations to the situation may end up changing the outcomes subjectively because they leak into the decision maker's experience.

The key idea in this argument is not that framing exerts effects on a decision, as is well known, but that its effects do not *stop* there. They can continue after the decision is made, when the results of the decision are being experienced. When these framing effects persist, the experience of the result of the decision can be consistent with the experience of the decision itself, and such consistency undermines the claim that the framing effect on the decision is irrational (see also Frisch & Clemen, 1994; Frisch & Jones 1993, for arguments that partly anticipate these arguments).

Belief in the importance of subjective experiences is widely shared by psychologists, economists, and many, many others. Indeed, much of Kahneman's more recent work (e.g. Kahneman, 2000) has been focused on subjective experience. He has developed the concept of *experienced utility*, which is based on a moment-by-moment evaluation of pleasure and pain, and he has used experienced utility as the standard for assessing the accuracy and rationality of people's predictions, decisions, and memories. But it is interesting to note here that despite Kahneman's attention to the nature of the experience of decisions over the last fifteen years, he has not combined this line of work

with the work he and Tversky did for 25 years on the determinants of the decisions themselves. It is as if heuristics, biases, and the dynamics of prospect theory exert their effects while a decision is being contemplated, but once the decision is made, experiencing the results of the decision will be “path independent.” The object of a decision will be experienced on its own, carrying no trace of how the decision was arrived at (but see Kahneman, 1994). It is this notion of path independence that Frisch’s findings challenge.

It is easy to accept the importance of subjective experience and still fail to apply this view in a thoroughgoing way when thinking about decision making. Path dependence is often hidden by a way of thinking that treats experiences as if they were caused directly by objects and episodes in the world, rather than as an interaction between the thing that is experienced and the person who is experiencing it. Gilbert and Ebert (2002) compared this “illusion of intrinsic satisfaction” to the perceptual illusion of direct access to the world, writing that “ordinary decision makers ignore the complexities of psychology and act instead as though their hedonic experiences were due entirely to the enduring intrinsic properties of their outcomes, as though the wonderfulness or awfulness they are experiencing was always there ‘in the outcome’ waiting to be experienced and none of these properties was altered or induced by the mere act of making the outcome their own” (p. 511). But this variant of “naïve realism” is mistaken: the value of an evening at a concert (in Examples 3 and 4), depends on how the person who is attending that concert construes and experiences the outcome (see also Gilbert, 2006).

There may be a sense in which decisions based on accurate predictions of their

objective consequences are made rationally, while those based on subjective consequences are non-rational, and those that result from false beliefs are irrational. However, this sense of “rationality” is not the only one, and maybe not even the most important (e.g. Parfit, 1984). Any decision-making process that helps people get better subjective results is not clearly contrary normative principles. Experienced utility matters, and procedural considerations that fail to take this fact into account are, at the very least, incomplete.

What I am calling “leakage” implies that there are limits on what formal principles of rationality can tell us, since there are surprisingly few cases where the formal principles apply in their strictest forms. Even seemingly inconsequential changes to the situation may leak into experience, affecting ones subjective outcomes and hence the reasonableness of different choices. Formal rules of rationality may allow researchers to draw important normative conclusions based on minimal, widely accepted structural claims about rationality. However, once the importance of subjective experiences and the prevalence of leakage are taken into account, it becomes clear that much more needs to be known before anything approaching a satisfactory theory of rationality is in hand. In particular, what is needed is a *substantive* theory of rationality—a theory that considers the content and not just the structure of decisions, and evaluates that content in light of the decision-maker’s goals and life as a whole. What is needed, in the words of Evans, Over, and Manktelow (1993), is a theory of rationality of *purpose*, to augment the formal theory of rationality of *process*.

I should make clear that I do not have a substantive theory of rationality to offer. What I can offer is a set of features that any adequate substantive theory should possess.

A substantive theory must consider the consequences of decisions, broadly construed. That is, it must consider short and long-term consequences, consequences to the self and to others, consequences that are central to the decision at hand and those that may be peripheral. It must also consider consequences of decisions for the character of the decision maker, since effects on character may have a significant impact on a host of future decisions. Research in behavioral decision making has in fact examined each of these classes of consequences, and has a good deal of interest to say about them. But virtually all of this previous work has been descriptive rather than normative. What I am suggesting is that a normative substantive theory is an essential complement to a normative formal theory if we are to be able to say anything significant about how well people make decisions.

An obvious choice for a substantive theory that incorporates the importance of subjective experience is the theory that the rational course of action is the one that maximizes experienced utility. Finding out what decisions will maximize experienced utility is a difficult empirical question, and, especially in particular cases, it may not be possible to have firm knowledge of the answer, but it is an interesting and important question that psychologists are in a good position to investigate (Gilbert, 2006; Kahneman, 2000).

However, the view that maximizing experienced utility is the rational thing to do involves an extensive normative commitment on a contentious philosophical question. It is one thing to say that experienced utility matters, but quite another to say that *only* experienced utility matters. And if experienced utility maximization is extended to situations where multiple people are involved, as a theory of morality in addition to

rationality, then it becomes more controversial still (Sen & Williams, 1988).

The importance of the subjective in evaluating decisions, coupled with the prevalence of leakage, seems to leave the normative theory of rational choice with little role to play in evaluating decisions. Does this mean that we should leave any effort to make normative claims behind? I do not think the leakage argument leaves us powerless, but it does demand a change of focus. The formal view of decision making often treats a decision as an isolated event. A person in a particular situation with a finite set of options (frequently two) can make a single choice from that set of options. Everything else is assumed to be fixed, including the decision maker's history, the shape of the paths that can be chosen from, and the future consequences of choosing one or another path.. However, this limited view of decision making neglects the ways in which the decision-making process is integrated into a person's life. Zooming out from this narrow view of decisions to a broader view can highlight problems with assessing rationality in isolated situations and also offer a path to making normative claims. Consider a man who is afraid of elevators. His isolated decision to trudge up the stairs is rational, in a sense, since it keeps him from experiencing the intense fear that would accompany any involvement with the elevator. However, it is strange to take this man's fear as a given, and to proceed from that assumption with narrow, means-ends reasoning. The rational thing for this man to do, either from a therapeutic point of view or from a broader consequentialist perspective, is to overcome his fear. Though trudging up the stairs *at this moment* may be the rational choice, trudging up stairs (exercise aside) as a life policy is not. It will interfere with too many other goals the person has.

A broad, substantive perspective on rationality thus permits an alternative to

either formal principles or experienced utility. The rationality of a decision can be assessed, at least in part, on the basis of the appropriateness of reactions to decisions once made. However, even if leakage into experience *is* irrational in a particular case, ignoring this leakage when making decisions is not the path to rationality. As Kahneman (2003) observes, “a theory of choice that completely ignores feelings such as the pain of losses and the regret of mistakes is not just descriptively unrealistic. It also leads to prescriptions that do not maximize the utility of outcomes as they are actually experienced” (p. 706). It is hardly rational for a man who fears elevators to choose to ride them in terror because he judges his fear to be irrational. In order to be rational he must eliminate his fear of elevators. If the problem is irrational feelings and related phenomena, then the solution must be for people to learn how to have the correct feelings. Irrational leakage into experience must be corrected with a systematic campaign of “leak plugging.” If people are costing themselves money by underinvesting in the stock market as a result of their tendency to evaluate their investments too frequently (Benartzi & Thaler, 1995), then the goal is to help people make money by teaching them not to be so concerned about the week-to-week or year-to-year performance of their investments. On this global view of rationality, which tries to integrate over multiple decisions and experiences, the emphasis is on making people more rational, sometimes by altering substantive characteristics of the person as decision maker and experiencer, and sometimes by altering formal procedures for evaluating options and making decisions. What rationality requires will depend on both the decision maker and the context within which the decision is made. The trick may be to value formal principles of rationality, but not take them too seriously.

Researchers can have a clear role in determining whether a formal principle has a good enough fit with the situation for the principle to be treated as applicable. Researchers need to be aware of the hidden psychological consequences—the leakage—of a decision process. If the leakage is negligible, then the formal principle will be useful. It can allow researchers to draw normative conclusions based on relatively uncontroversial formal claims about rationality. However, if leakage has a sizeable influence, then researchers have a good reason for denying the application of the formal principle.

Is there anything to be said in a general way about when formal principles are most likely to be powerfully relevant or to come closest to applying)? In some cases, formal principles may apply well to a situation because of the importance of the decision to the welfare of others or to one's own future welfare. When deciding between radiation and surgery as a treatment for cancer, the obvious target factor—the influence of the treatment on survival rates—is so important that it seems wrong for decisions to change when the survival rates are reframed as mortality rates (McNeil, Pauker, Sox & Tversky, 1982). And in the Asian disease example, letting the frame influence a decision seems frivolous because other people's lives are at stake. Consistent with this speculation that leakage should be less of an issue when one is making decisions for others, there is evidence that people are sometimes more resistant to framing effects when they are making decisions for other people (Krishnamurthy & Kumar, 2002; see also Hsee & Weber 1997, who attribute a similar result to the role of emotions in decision making). A judge who is deciding which parent should be granted custody of a child may make a different choice if the question is framed as which parent should be denied custody

(Shafir, Simonson, & Tversky, 1993). But the important consequences here are for the child and the parents, so the effects of the decision frame on the judge's subjective experience should be irrelevant. Indeed, the judge's job demands that his or her subjective experience of the decision be irrelevant to the making of the decision, just as a doctor guiding a patient through the choice of surgery or chemotherapy should be sensitive to the patient's experienced utility, but insensitive to his or her own. These examples combine highly important target factors with a separation between the decision maker and the people who experience the consequence of the decision, as does the Asian disease scenario, which also multiplies the importance of the target factor by putting hundreds of people's lives at stake.

Neutralizing leakage is one reason why it may often be advantageous to separate the "principal" from the "agent" in decision contexts, even when the agent knows little more about the domain in question than the principal. By blocking leakage from the decision to the experience, principal-agent separation can escape the tradeoff between "objectively" good decisions and "subjectively" good decisions, freeing the experiencer—the principal—from negative leakage and allowing the decision maker—the agent—to investigate the merits of the options without worrying about any potential leakage from that investigation.

To conclude this section of the paper, although formal principles of decision making like invariance often seem compelling, there is only a narrow set of cases in which they apply in their strictest form. In most cases, there is too much going on psychologically—too much leakage from decision to experience—for these principles to capture all the relevant features of the situation. It is possible to apply a formal principle

in those cases where the abstract structure of the principle does not strictly or fully apply, but it is only normatively appropriate to do so if the substantive claim is correct that the amount of leakage is negligible. A substantive approach can help assess whether leakage effects are minor enough to neglect or important enough to override the formal principle. Formal principles may also be useful as prescriptive rules that help people act rationally rather than as normative ones that define rationality (Baron, 1986). But a broader approach is necessary in order to evaluate the effects of adopting normative principles as prescriptive rules, as well as to evaluate more explicit advocacy, training, or intervention designed to promote the normative principles. The world and its human inhabitants are too complex for us to be able to arrive at satisfying comprehensive answers to many of our most significant questions. But because the stakes are so high, it is important to try to address these big normative questions and to work to develop normative frameworks for decision making alongside of our descriptive theories.

#### Non-Monotonicity of “Rationality”

“Both excessive and defective exercise destroys the strength, and similarly drink or food which is above or below a certain amount destroys the health, while that which is proportionate both produces and increases and preserves it. So too is it, then, in the case of temperance and courage and the other virtues. For the man who flies from and fears everything and does not stand his ground against anything becomes a coward, and the man who fears nothing at all but goes to meet every danger becomes rash; and similarly the man who indulges in every pleasure and abstains from

none becomes self-indulgent, while the man who shuns every pleasure, as boors do, becomes in a way insensible; temperance and courage, then, are destroyed by excess and defect, and preserved by the mean.”

– Aristotle (1999, p. 22)

The previous section of the paper suggested that whereas formal principles of rational choice are a good thing, there can be too much of a good thing. Similarly, as the literature on choice overload suggests, though choice is a good thing, there can be too much of that as well. These sorts of findings and arguments led Adam Grant and me to ask whether this non-monotonicity was anomalous or common (Grant & Schwartz, 2011). After reviewing several different areas of research, we concluded that non-monotonicity was pervasive.

This is something that Aristotle understood. When he explored what makes for a happy and successful life, Aristotle (1999) concluded that happiness and success are a function of cultivating virtues that lie at the mean between the extremes of deficiency and excess (Nussbaum, 1995, 2004). From Aristotle’s viewpoint, the effects of virtues on well being takes the shape of an inverted—U. For example, in the domain of self-presentation, honesty is the mean between the deficiency of self-deprecation and the excess of boastfulness. In the domain of pleasing others, friendliness is the mean between the deficiency of quarrelsomeness and the excess of ingratiating. What insights does psychology offer about the effects of deficiency and excess on success and well being?

Building on the notion of the Aristotelian mean, there is good reason to believe that many character traits and environmental influences have non-monotonic effects.

Suedfeld (1969) referred to this principle as the “ubiquitous U”: across many domains of psychology, one finds that X increases Y to a point, and then it decreases Y. Common examples include the Yerkes-Dodson law of motivation (1908) and classic theories of optimal arousal (Eysenck, 1967; Smith, 1983). My purpose here is to draw attention to what may be a fundamental and ubiquitous psychological principle: there is no such thing as an unmitigated good, including the adherence to formal principles of rational choice. Many “positive” traits, states, and experiences have costs that at high levels may begin to outweigh their benefits, creating the non-monotonicity of an inverted-U.

Grant and Schwartz (2011) discuss many different examples of inverted-U type effects, and I will highlight only a handful of them here. Bunderson and Sutcliffe (2003) studied the learning orientations of management teams in a Fortune 100 consumer products company. Past research had shown a positive relationship between team learning orientation and business performance. Bunderson and Sutcliffe (2003, p. 554) proposed that strong learning orientations can be “inefficient or even dysfunctional,” as a focus on learning can distract attention away from performance results, lead members to waste resources on low-probability experiments, and create divergent responsibilities and discoveries that make it difficult to assimilate and disseminate knowledge gained. And they found a curvilinear, non-monotonic relationship between team learning orientation and business performance.

Some individuals seek out complex jobs, which provide opportunities for learning, challenges, skill development, and growth (e.g., Holland, 1996; Judge & Bretz, 1992). Although researchers initially assumed psychological benefits of job complexity, more recent evidence suggests that high complexity exacerbates stress, burnout, and

dissatisfaction. A number of studies have revealed inverted-U-shaped relationships between job complexity and well being, such that well being is highest at moderate levels of complexity and lower at high levels of complexity (Champoux, 1982; Janssen, 2001; Xie & Johns, 1995).

People differ in their willingness to persist in the face of challenges, and the presumption has been that the more persistent people are, the more they will achieve. Although research has demonstrated that persistent practice facilitates the development of expertise (Ericsson & Charness, 1994), there is also evidence that practice can be most productive in moderation. Langer and Imber (1979) found that individuals who engaged in moderate practice prior to a task made fewer errors than those who engaged in no practice or extensive practice. Although some practice facilitated performance, high levels of practice led to overlearning and inflexibility, preventing individuals from effectively accessing and improving their skills (Langer & Imber, 1979). Consistent with this experimental evidence, in a study of NBA basketball teams, Berman, Down, and Hill (2002) found that the longer teams practiced and played together, the more games they won. This relationship was mediated by better coordination between players, in the form of assists. However, this benefit of persistent practice dissipated after approximately four years, at which point experience appeared to result in overconfidence, complacency, and routine rigidity. Corroborating these findings, meta-analytic evidence suggests that in many jobs, the relationship between experience and performance takes the form of an inverted-U (Sturman, 2003).

Further, although extensive research has linked persistence in the face of failure to higher performance in a variety of tasks (e.g., Dweck, 2006; Grant, Campbell, Chen,

Cottone, Lapedis, & Lee, 2007), there is evidence that very high levels of persistence can undermine performance. Moon (2001) demonstrated that highly persistent, conscientious individuals who strongly value achievement are more likely to escalate their commitment to failing courses of action, investing time, money, and resources in losing endeavors. Such dysfunctionally high levels of persistence may undermine psychological and physical well being by preventing individuals from disengaging from goals at appropriate times (Miller & Wrosch, 2007; Wrosch, Scheier, Carver, & Schulz, 2007).

Research has also demonstrated inverted-U-shaped relationships between optimism and performance (Brown & Marshall, 2001). At moderate levels, optimism provides confidence and increases planning, but very high optimism leads to inadequate preparation and the underestimation of risks. Similarly, research shows that high levels of self-efficacy constitute overconfidence, which can result in groupthink (Whyte, 1998), persistence with failing strategies (Whyte & Saks, 2007), less time and energy invested in learning and planning (Vancouver & Kendall, 2006), and poor performance (Dunning, Heath, & Suls, 2004; Hayward & Hambrick, 1997; Stone, 1994; Vancouver, Thompson, Tischner, & Putka, 2002). Very high levels of optimism also appear to be costly because they encourage riskier health behaviors and high expectations that are difficult to meet (Milam, Richardson, Marks, Kemper, & McCutchan, 2004). Baumeister, Campbell, Krueger, and Vohs (2003) summarize similar evidence on self-esteem, suggesting that at very high levels, self-esteem can be costly to performance, interpersonal relationships, and health.

Flynn (2003) examined the possibility of non-monotonic effects of generosity on performance. Using data on helping exchanges between professional engineers, he found

an inverted-U-shaped relationship between generosity (giving more than one received) and productivity (quality, quantity, and efficiency of work completed relative to managers' goals). At moderate levels, generosity allows employees to gain status and call in favors from colleagues, but at high levels, generosity consumes time, energy, and other finite resources that make it more difficult to complete one's own work.

Although psychologists have traditionally demonstrated that empathy increases prosocial behavior (for a review, see Batson, 1998), there is evidence that very high levels of empathy can be emotionally aversive and undermine prosocial behavior. Eisenberg (2000, p. 674) summarized research on the phenomenon "empathic overarousal," in which a strong experience of empathy cultivates feelings of distress, which have the boomerang effect of distracting attention away from others and toward managing one's own aversive feelings (see also Hoffman, 1982, and Eisenberg, Fabes, Murphy, Maszk, Smith, O'Boyle, & Suh, 1994). Research also suggests that empathy runs the risk of undermining task performance. High levels of empathy can cloud judgment, leading to self-sacrificing behaviors that benefit others at the expense of achieving one's own goals (Galinsky, Maddux, Gilin, & White, 2008), or sometimes, even, fail to benefit others. In addition, research demonstrates that high empathy can encourage unethical behaviors that help the targets of empathy but violate principles of fairness and justice (Batson, Klein, Highberger, & Shaw, 1995; Gino & Pierce, 2009). *The Vices of Virtues: Toward a General Explanation.* The above evidence suggests that many characteristics that are linked to well being and strong performance can, at high levels, undermine the outcomes they are intended to promote. This support for the inverted-U highlights the value of gaining a more systematic understanding of the general

mechanisms that govern the non-monotonic effects. Grant and I (2011) suggested three general processes by which non-monotonicity may arise. One general mechanism concerns potential conflicts between different processes or characteristics. These types of conflicts are apparent in some of the studies reviewed above, such as in the cases of high levels of learning compromising achievement, practice reducing cognitive and behavioral flexibility, generosity limiting achievement and psychological well being, and empathy conflicting with justice. This kind of mechanism is consistent with the Aristotelian perspective, which views virtues as interdependent goods that must be cultivated in balance. As Schwartz and Sharpe (2006, p. 380) argued, “Virtues and strengths should not be treated in isolation from each other; they are not effective, in general, if exercised independently... nurturing a single signature strength can produce deformations of character.”

A second general mechanism concerns differences in shapes, slopes, and thresholds for positive and negative effects. For example, consider the sensory delight of eating delicious food. As you continue to stuff yourself at a buffet table that offers a copious supply of many treats, two things happen: first, there is diminishing marginal benefit to the delicious tastes, whether through adaptation or some other process. Second, you start to feel full—uncomfortably so. Early in your meal, there are no negative effects. But as you keep eating, a threshold is crossed, and negative effects start to appear. In addition, unlike the pleasurable effects, which diminish, the negative ones may escalate as they grow in magnitude. Indeed, Coombs and Avrunin (1977) suggested that as a general matter, “good things satiate; bad things escalate.” The combination of diminishing marginal utility of positive influences and escalating disutility of negative

influences whose threshold of activation differs from the positive effect and/or whose strength of activation increases non-linearly can produce the inverted-U. Schwartz (2009) has suggested that exactly this sort of mechanism underlies the choice overload effect.

A third general mechanism concerns the possibility that there is a single set of effects that is in and of itself non-monotonic. For example, the Yerkes-Dodson Law (1908) suggests that increases in motivation enhance effort and narrow attention. This improves performance in simple tasks, but impairs performance in complex tasks, for which narrowed attention can be a liability. The well-documented phenomenon of stereotype threat (e.g., Steele, 1997; Steele & Aronson, 1995) can be understood as resulting in part from Yerkes-Dodson type effects, as demonstrated by O'Brien and Crandall (2003). In their study, male and female students were told they would complete a series of math tests that either have shown gender differences (stereotype threat for women) or have not shown gender differences (control). The difficulty of the tests was varied so that half of the participants completed an easy set of math problems and the other half completed a difficult set of problems. Stereotype threat improved performance for women on the easy set of problems but harmed performance on the difficult set. Men were unaffected by the stereotype threat manipulation. What this example demonstrates is non-monotonicity that results directly from the non-monotonic effects of the variable under investigation (narrowing of attention), and not from either conflicting virtues or independent positive and negative effects that have different thresholds and different functional forms.

## Conclusion: The Aristotelian Mean, Inverted-U's, Satisficing, Leaks, and Rationality

The preceding section of the paper suggested that as a general matter, there can be too much of a good thing, and that rational people must be able to judge from context how much is enough. The section on robust satisficing argued that there are contexts in which it is normatively appropriate to seek to maximize utility and contexts in which it is not. In making decisions, people must judge what sort of context they are facing. And the section on leaky rationality argued that there are contexts in which people should pay attention to leaks and contexts in which they should not—contexts in which they should consider the effects of their potential decisions on their lives as a whole and the lives of other people, and contexts in which it is appropriate to indulge their pursuit of experienced utility. The key word in each of these arguments is “judgment,” and I suggest, with Aristotle, that rationality is good judgment, or as he put it, “practical wisdom” (see Schwartz & Sharpe, 2011). I think that practical wisdom is the hallmark of rationality. And as Schwartz and Sharpe argue, it is, at its core, the substitution of judgment for rules, including the rules of rational choice theory (see also Baltes & Staudinger, 2000; Schwartz & Sharpe, 2006; Sternberg, 1998). Practical wisdom is a substantive, not a formal characteristic, and its value is revealed in substantive life outcomes, not in logical consistency, transitivity, or invariance.

## References

- Aristotle. (1999). *Nicomachean ethics* (W. D. Ross, Trans.) Kitchener: Batoche Books.
- Baltes, P. B., & Staudinger, U. M. (2000). Wisdom: A metaheuristic (pragmatic) to orchestrate mind and virtue toward excellence. *American Psychologist*, *55*, 122-136.
- Baron, J. (1986). Tradeoffs among reasons for action. *Journal for the Theory of Social Behavior*, *16*, 173-195.
- Baron, J. (2004). Normative models of judgment and decision making. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp.19-36). Malden, MA.: Blackwell.
- Baron, J. (2008). *Thinking and deciding* (4th Ed.). New York: Cambridge.
- Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology, volume 2, 4<sup>th</sup> ed.* (pp. 282-316). New York: McGraw-Hill.
- Batson, C. D., Klein, T. R., Highberger, L., & Shaw, L. L. (1995). Immorality from empathy-induced altruism: When compassion and justice conflict. *Journal of Personality and Social Psychology*, *68*, 1042-1054.
- Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, *4*, 1-44.
- Benartzi, S. & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics*, *110*, 73-92.

- Ben-Haim, Y. (2006). *Info-gap decision theory: Decisions under severe uncertainty* (2nd edition). London: Academic Press.
- Ben-Haim, Y. (2011). *Info-Gap Economics: An Operational Introduction*. New York: Palgrave-Macmillan.
- Berman, S. L., Down, J., & Hill, C. W. L. (2002). Tacit knowledge as a source of competitive advantage in the National Basketball Association. *Academy of Management Journal*, 45, 13-31.
- Bolino, M. C., & Turnley, W. H. (2005). The personal costs of citizenship behavior: The relationship between individual initiative and role overload, job stress, and work-family conflict. *Journal of Applied Psychology*, 90, 740-748.
- Botti, S., & Iyengar, S.S. (2004). The psychological pleasure and pain of choosing: When people prefer choosing at the cost of subsequent outcome satisfaction. *Journal of Personality and Social Psychology*, 87, 312-326.
- Botti, S. & Iyengar, S.S. (2006). The dark side of choice: When choice impairs social welfare. *Journal of Public Policy and Marketing*, 25, 24-38.
- Brown, R. V. (1993). Impersonal probability as an ideal assessment based on accessible evidence: A viable and practical construct? *Journal of Risk and Uncertainty*, 7, 215–235.
- Brown, J. D., & Marshall, M. A. (2001). Great expectations: Optimism and pessimism in achievement settings. In E. C. Chang (Ed.), *Optimism & pessimism: Implications for theory, research, and practice* (pp. 239-255). Washington, DC: American Psychological Association.
- Bunderson, J. S., & Sutcliffe, K. M. (2003). Management team learning orientation and

- business unit performance. *Journal of Applied Psychology*, 88, 552-560.
- Champoux, J. E. (1992). A multivariate analysis of curvilinear relationships among job scope, work context satisfactions, and affective outcomes. *Human Relations*, 45, 87-111.
- Coombs, C. H. & Avrunin, G.S. (1977). Single peaked functions and the theory of preference. *Psychological Review*, 84, 216-30.
- Dar-Nimrod, I., Rawn, C. D., Lehman, D. R., & Schwartz, B. (2009). The maximization paradox: The costs of seeking alternatives. *Personality and Individual Differences*, 46, 631-635.
- Diab, D. L., Gillespie, M. A., & Highhouse, S. (2008). Are maximizers really unhappy? The measurement of maximizing tendency. *Judgment and Decision Making*, 3, 364-370.
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5, 69-106.
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.
- Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology*, 51, 665-697.
- Eisenberg, N., Fabes, R. A., Murphy, B., Karbon, M., Maszk, P., Smith, M., O'Boyle, C., & Suh, K. (1994). The relations of emotionality and regulation to dispositional and situational empathy-related responding. *Journal of Personality and Social Psychology*, 66, 776-797.

- Ellsberg, D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics*, 75, 643–669.
- Ericsson, K. A. & Charness, N. 1994. Expert performance: Its structures and acquisition. *American Psychologist*, 49, 725-747.
- Evans, J. St. B. t., Over, D. E., & Manktelow, K. I. (1993). Reasoning, decision making and rationality. *Cognition*, 49, 165-187.
- Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, IL: Charles C. Thomas.
- Flynn, F. J. (2003). How much should I give and how often? The effects of generosity and frequency of favor exchange on social status and productivity. *Academy of Management Journal*, 46, 539-553.
- Frisch, D. (1993). Reasons for framing effects. *Organizational Behavior and Human Decision Processes*, 54, 399-429.
- Frisch, D. & Clemen, R. T. (1994). Beyond expected utility: Rethinking behavioral decision research, *Psychological Bulletin*, 116, 46-54.
- Frisch, D. & Jones, S. K. (1993). Assessing the accuracy of decisions. *Theory and Psychology*, 3, 115-135.
- Galinsky, A. D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. *Psychological Science*, 19, 378-384.
- Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. New York: Viking.
- Gilbert, D. T. (2006). *Stumbling on happiness*. New York: Knopf.
- Gilbert, D. T. & Ebert, J. E. (2002). Decisions and revisions: The affective forecasting of

- changeable outcomes. *Journal of Personality and Social Psychology*, 82, 503-514.
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.) (2002)., *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.
- Grant, A. M., Campbell, E. M., Chen, G., Cottone, K., Lapedis, D., & Lee, K. (2007). Impact and the art of motivation maintenance: The effects of contact with beneficiaries on persistence behavior. *Organizational Behavior and Human Decision Processes*, 103, 53-67.
- Grant, A.M. & Schwartz, B. (2011). Too much of a good thing: The challenge and opportunity of the inverted-U. *Perspectives on Psychological Science*, 6, 61-76.
- Hanoch, Y., Wood, S., Barnes, A., Liu, P., & Rice, T. (2011). Choosing the right medicare prescription drug plan: The effect of age, strategy selection, and choice set size. *Health Psychology*, 30, 719-727.
- Hayward, M. L. A., & Hambrick, D. C. (1997). Explaining the premiums paid for large acquisitions: Evidence of CEO hubris. *Administrative Science Quarterly*, 42, 103–127.
- Hoffman, M. L. (1982). Development of prosocial motivation: empathy and guilt. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 281-313). New York: Academic Press.
- Holland, J. L. (1996). Exploring careers with a typology: What we have learned and some new directions. *American Psychologist*, 51, 397-406.
- Hsee, C. K. & Weber, E. U. (1997). A fundamental prediction error: Self–other discrepancies in risk preference. *Journal of Experimental Psychology: General*, 126, 45–53.

- Iyengar, S. S. (2010). *The art of choosing*. New York: Twelve.
- Iyengar, S. S., Jiang, W., & Huberman, G. (2004). How much choice is too much: Determinants of individual contributions in 401K retirement plans. In O.S. Mitchell & S. Utkus (Eds.) *Pension design and structure: New lessons from behavioral finance* (pp. 83-95). Oxford: Oxford University Press.
- Iyengar, S. S. & Lepper, M. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79, 995-1006.
- Iyengar, S.S., Wells, R. E. , & Schwartz, B. (2006). Doing better but feeling worse: Looking for the “best” job undermines satisfaction. *Psychological Science*, 17, 143-150.
- Judge, T. A., & Bretz, R. D. Jr. (1992). Effects of work values on job choice decisions. *Journal of Applied Psychology*, 77, 261-271.
- Kahneman, D. (1994). New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics*, 150, 18-36.
- Kahneman, D. (2000). Experienced utility and objective happiness: A moment-based approach. In D. Kahneman, & A. Tversky (Eds.), *Choices, Values, and Frames* (pp. 673-692). New York: Cambridge University Press.
- Kahneman, D. (2003). A perspective on judgment and choice. *American Psychologist*, 58, 697-720.
- Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.
- Kahneman, D. and Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. pp. 49-81 in T. Gilovich, D. Griffin, & D.

- Kahneman [eds]. *Heuristics & Biases: The Psychology of Intuitive Judgment* (pp.49-81). New York. Cambridge University Press.
- Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under uncertainty. *Econometrica*, 47, 263-291.
- Kahneman, D. & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39, 341-350.
- Kahneman, D. & Tversky, A. (Eds.) (2000). *Choices, values, and frames*. New York: Cambridge University Press.
- Keeney, R. L. & Raiffa, H. (1993). *Decisions with multiple objectives—preferences and value tradeoffs*. New York: Cambridge University Press.
- Keys, D. J. & Schwartz, B. (2007). “Leaky” rationality: How research on behavioral decision making challenges normative standards of rationality. *Perspectives on Psychological Science*, 2, 162-180.
- Knight, F.H. (1921). *Risk, uncertainty and profit*. Originally published by Hart, Schaffner and Marx. Re-issued 1965, New York: Harper Torchbooks.
- Krishnamurthy, P. & Kumar, P. (2002). Self-other discrepancies in waiting time decisions. *Organizational Behavior and Human Decision Processes*, 87, 207-226.
- Langer, E. J., & Imber, L. G. (1979). When practice makes imperfect: Debilitating effects of overlearning. *Journal of Personality and Social Psychology*, 37, 2014-2024.
- Levin, I. P. & Gaeth, G. J. (1988). How consumers are affected by the framing of attribute information before and after consuming the product. *Journal of Consumer Research*, 15, 374-378.
- McNeil, B. J., Pauker, S. G., Sox, H. C. Jr., & Tversky, A. (1982). On the elicitation of

preferences for alternative therapies. *New England Journal of Medicine*, 306, 1259-1262.

Milam, J. E., Richardson, J. L., Marks, G., Kemper, C. A., & McCutchan, A. J. (2004). The roles of dispositional optimism and pessimism in HIV disease progression. *Psychology and Health*, 19, 167-181.

Miller, G. E., & Wrosch, C. (2007). You've gotta know when to fold 'em: Goal disengagement and systematic inflammation in adolescence. *Psychological Science*, 18, 773-777.

Nisbett, R. E. and Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.

Nussbaum, M. C. (1995). *Poetic justice*. Boston: Beacon Press.

Nussbaum, M. C. (2001). *Upheavals of thought*. New York: Cambridge University Press.

Nussbaum, M. C. (2004). *Positive psychology and ancient Greek virtue ethics*. Paper presented at The Philosophical History of the Virtues conference, University of Pennsylvania, September.

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, 29, 782-789.

Over, D. (2004). Rationality and the normative/descriptive distinction. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp.3-18). Malden, MA.: Blackwell.

Parfit, D. (1984). *Reasons and persons*. Oxford: Clarendon Press.

Polman, E. (2010). Why are maximizers less happy than satisficers? Because they

- maximize positive and negative outcomes. *Journal of Behavioral Decision Making*, 23, 179-190.
- Sanford, A. J., Fay, N., Stewart, A., & Moxey, L. (2002). Perspective in statements of quantity, with implications for consumer psychology. *Psychological Science*, 13, 130-134.
- Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.
- Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2009). What moderates the too-much-choice effect? *Psychology & Marketing*, 26, 229-253.
- Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, 37, 409-425.
- Schwartz, B. (2004). *The paradox of choice: Why more is less*. New York: Ecco Press.
- Schwartz, B. (2009). Be careful what you wish for: The dark side of freedom. In R. M. Arkin, K. C. Oleson, & P. J. Carroll (Eds.), *Handbook of the uncertain self: Perspectives from social and personality psychology* (pp. 62-77). New York: Psychology Press.
- Schwartz, B., Ben-Haim, Y., & Dacso, C. (2011). What makes a good decision? Robust satisficing as a normative standard of rational decision making. *Journal for the Theory of Social Behaviour*, 41, 209-227.
- Schwartz, B., & Sharpe, K. (2006). Practical wisdom: Aristotle meets positive psychology. *Journal of Happiness Studies*, 7, 377-395.
- Schwartz, B. & Sharpe, K. (2011). *Practical wisdom: The Right Way to Do the Right Thing*. New York: Riverhead.

- Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D.R. (2002). Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology, 83*, 1178-1197.
- Seligman, M.E.P. (2002). *Authentic happiness*. New York: Free Press.
- Sen, A. & Williams, B. (1988). *Utilitarianism and beyond*. New York: Cambridge University Press.
- Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition, 49*, 11-36.
- Shiv, B., Carmon, Z., & Ariely, D. (2005). Placebo effects of marketing actions: Consumers may get what they pay for. *Journal of Marketing Research, 42*, 383-393.
- Simon, H.A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 59*, 99-118.
- Simon, H. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129-138.
- Simon, H. (1957). *Models of man, social and rational*. New York: Wiley.
- Smith, B. D. (1983). Extraversion and electrodermal activity: Arousability and the inverted-U. *Personality and Individual Differences, 4*, 411-419.
- Sparks, E. A., Ehrlinger, J., & Eibach, R. P. (2012). Failing to commit: Maximizers avoid commitment in a way that contributes to reduced satisfaction. *Personality and Individual Differences, 52*, 72-77.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*: 613-629.
- Steele C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test

- performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797-811.
- Sternberg, R. J. (1998). A balance theory of wisdom. *Review of General Psychology*, 2, 347-365.
- Stone, D. N. (1994). Overconfidence in initial self-efficacy judgements: Effects on decision processes and performance. *Organizational Behavior and Human Decision Processes*, 59, 452-474.
- Sturman, M. C. (2003). Searching for the inverted U-shaped relationship between time and performance: Meta-analyses of the experience/performance, tenure/performance, and age/performance relationships. *Journal of Management*, 29, 609-640.
- Suedfeld, P. (1969). The delectable "D" and the ubiquitous "U " In D. G. Forgas & J Levin, Arousal theory NASA-University of Vermont Symposium, Burlington, VT, September.
- Sunstein, C. R., Kahneman, D., Schkade, D. & Ritov, I. (2002). Predictably incoherent judgments. *Stanford Law Review*, 54, 1153-1216.
- Thaler, R. H. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1, 39-60.
- Thaler, R.H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, 12, 183-206.
- Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453-458.
- Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to

- motivation and performance in a learning context. *Journal of Applied Psychology*, *91*, 1146-1153.
- Vancouver, J. B., Thompson, C. M., Tischner, E. C., & Putka, D. J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of Applied Psychology*, *87*, 506-516.
- Whyte, G. (1998). Recasting Janis's groupthink model: The key role of collective efficacy in decision fiascoes. *Organizational Behavior and Human Decision Processes*, *73*, 185-209.
- Whyte, G., & Saks, A. M. (2007). The effects of self-efficacy on behavior in escalation situations. *Human Performance*, *20*, 23-42.
- Wilson, T. D., & Gilbert, D. T. 2005. Affective forecasting: Knowing what to want. *Current Directions in Psychological Science*, *14*: 131-134.
- Wrosch, C., Scheier, M. F., Carver, C. S., & Schulz, R. (2007). The importance of goal disengagement in adaptive self-regulation: When giving up is beneficial. *Self and Identity*, *2*, 1-20.
- Xie, J. L., & Johns, G. (1995). Job scope and stress: Can job scope be too high? *Academy of Management Journal*, *38*, 1288-1309.
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*, 459-482.

