

Running head: BROADENING BEHAVIORAL DECISION RESEARCH

Broadening Behavioral Decision Research:  
Multiple Levels of Cognitive Processing

Douglas L. Medin

Northwestern University

Max H. Bazerman

Northwestern University and Harvard University

### Abstract

The area of behavioral decision research, specifically the work on heuristics and biases, has had a tremendous influence on basic research, applied research, and application over the last twenty-five years. Its unique juxtaposition against economics has provided important benefits, but at the cost of leaving it disconnected from too much of psychology. This paper explores an expanded definition of behavioral decision research through the consideration of multiple levels of cognitive processing. Rather than being limited to how decision-makers depart from optimality, we offer a broader analysis of how decision-makers define the decision problem and link decisions to goals, as well as a more detailed focus on processes associated with implementing decisions.

Over the past few decades the area of cognitive psychology has grown dramatically, social and developmental psychology have moved strongly in a cognitive direction, and behavioral decision research (BDR) has emerged as a new area of psychology. BDR is unique among psychological subfields in the impact that it has had on research outside of psychology - including its impact on economics, finance, public policy, law, medicine, marketing, organizational behavior, and negotiation. Unfortunately, BDR has also moved further away from many core areas of psychology, limiting its theoretical development and its integration with advances made in allied areas.

Our central thesis is that the most well known part of BDR, the heuristics and biases approach, has been overly constrained by a focus on how people make mistakes at the point of decision. Research on heuristics and biases has implicitly assumed that the goal is known and that the details of implementing decisions are not part of the problem. The prescriptive goal is optimality defined in terms of the behavior of rational agents and, from that perspective, the heuristic and biases literature has yielded a fascinating catalog of human decision errors that is important for both theoretical and practical reasons. A drawback with this orientation, however, is that it tends to define human decision making by what it is not. As a consequence, it does not provide an effective framework for the detailed study of decision makers' (often multiple) goals that serve to define a decision context nor the linkages between goals and processing issues associated with implementing decisions. We seek to add these components to decision making analyses, without losing the key strengths of the heuristics and biases approach.

This paper overviews the historical development of BDR, highlights a number of challenges to BDR, examines the potential of considering multiple levels of analysis to organize

new directions for BDR, illustrates the potential for a more psycho-central perspective to decision making within this framework, and provides examples of recent and ongoing research across levels. Collectively, our aim is to build on the strengths of BDR by broadening the framework used to understand how we make decisions. In pursuing this strategy we are building on related suggestions by Montgomery (1983), Kahneman (1991), Frisch and Clemen (1994), Hilton (1995) and chapters in the edited volume by Busemeyer, Hastie, and Medin (1995).

### Behavioral Decision Research: A Brief History

The most unique difference between the heuristics and biases approach to BDR and other areas of psychology is its use of economic rationality as the backdrop for descriptions of actual decisions. Economics has traditionally assumed that individuals will act rationally in the pursuit of their own interests. Echoing this sentiment, Anderson (1991) has argued that the best way to develop models of cognition is to analyze what an agent's goals are and then determine a mechanism that would yield optimal behavior.

By way of contrast, research on heuristics and biases has been characterized by a delineation of the systematic ways in which people deviate from optimality or rationality (Kahneman and Tversky, 1973, 1979; Dawes, 1997). Consistent with Simon (1957, March and Simon, 1958), individuals are presumed to attempt to act rationally, but to be bounded in their ability to achieve rationality. The four decades since Simon's original statement on bounded rationality have seen the articulation of the multitude of specific ways in which we fall short of rational behavior. The result is that researchers can predict, a priori, how people will make decisions that are inconsistent, inefficient, and based on normatively irrelevant information.

By any account, this research has been enormously successful with respect to both theory

and application. Generalizations from laboratory studies have held up well when tested in real world contexts (Northcraft and Neale; 1987; Camerer, 1995; Dawes, 1997; Bazerman, 1998; Gilovich, 1992). As we shall see, theoretical analyses are becoming increasingly sophisticated. Still, we do not think it would be too much of an exaggeration to say that BDR currently suffers from something of a malaise (see Goldstein and Weber, 1995) or perhaps a sense that something is missing. Lest we be accused of undue pessimism, one might equally say that BDR has succeeded so well that the traditional framework is increasingly less apt as a description of ongoing research. The category label, “studies showing people are not optimal decision makers,” is far less coherent than it once was, because many novel lines of research are being developed and there is no clear central tendency. We see current trends as re-examining foundational assumptions of BDR. Before turning to these considerations we take a moment to place our arguments in context.

The essential argument of much of BDR, consistent with Simon (1957), is that people rely on simplifying strategies, or cognitive heuristics. While these heuristics are frequently useful short cuts, they also lead us astray (Tversky and Kahneman, 1974). People fail to make optimal choices, because the optimal choice has not been examined or has been eliminated through heuristic processing. By providing a simple method for making decisions, heuristics produce pretty good decisions most of the time. However, these heuristics also lead to biases. It is the systematic and predictable nature of these biases that has made them so intriguing to researchers.

People use many heuristics. Occupations develop their own rules of thumb. Mortgage bankers believe that “people can only afford 35% of their income for housing”. This seems like a fairly rigid rule, when mortgage applicants differ in their income from \$20,000 to \$100,000 to \$400,000 - all these people still only eat one dinner per night. But this heuristic makes processing

an application a much easier task. Similarly, the difference between a p-statistic of .04 and .05 is treated as a smaller difference by reviewers than the difference between .045 and .055. We have a heuristic of " $p < .05$ " that makes scientific assessment easier. However, the heuristics that are central to the field of BDR are generally those that are less specialized and have widespread application. We turn now to a brief considerations of some examples.

Tversky and Kahneman (1971) showed that people expect future observations to be like past ones - even when they have little data on which to make such predictions. For example, Tversky and Kahneman argued that people ignore sample size, and apply a "law of small numbers" to make predictions in situations where, normatively speaking, the "law of large numbers" is appropriate. More generally, people are often insensitive to sample size. Kahneman and Tversky (1972) later subsumed "the law of small number" within the broader heuristic of representativeness. Representativeness is the tendency for people to assess the likelihood of an event based on the similarity of that occurrence to their stereotypes of similar occurrences (Kahneman and Tversky, 1972).

Availability and anchoring and adjustment were then added as additional general heuristics. Availability is the tendency to assess the frequency, probability, or likely causes of an event by the degree to which instances or occurrences of that event are readily "available" in memory (Tversky and Kahneman, 1973). For example, an event that is vivid, easily imagined, and specific will be more "available" than will an event that is unemotional in nature, bland, difficult to imagine, or vague. Anchoring and adjustment refers to the tendency to assess quantities by starting from an initial value and adjusting to yield a final decision (Tversky and Kahneman, 1974). It is well-documented that people are overly influenced by anchors, even arbitrary ones. The initial value may be suggested from historical precedent (which may no longer be

appropriate), from the way in which a problem is presented, or from random information. Even a trivial factor can have a profound effect on our decision if it serves as a starting point from which we make adjustments (Dawes, 1988).

Over the last quarter century or so many systematic and generalizable biases have been identified as a result of our heuristic processing, and the list continues to grow. We will not discuss these effects in detail here. A number of reviews are available (Dawes, 1997; Bazerman, 1998; Kahneman, Slovic, and Tversky, 1982; Thompson, 1998). We note, however, that it is the demonstration of these biases as departures from rational thought that has created the widely diffused interest in BDR. Doctors, lawyers, consultants, accountants, and diplomats want to make rational decisions, and are intrigued by evidence that they may not. More importantly, the systematic nature of our biases allows for the potential to specify the changes that are necessary in human cognition to come closer to rationality. There is no doubt that this juxtaposition against rationality has led to BDR being a new lens of analysis for the many applied areas listed above.

Fischhoff (in press) argues that the broad impact of BDR can be traced to four converging factors. First, initial demonstrations have proven to be quite robust, replicable, and generalizable. Second, the effects have been presented in fascinating ways that capture the imagination of individuals well beyond the scientific literature. Third, BDR fits the cognitive revolution's focus on tracing failures to unintended side effects of adaptive processes. Finally, heuristics and biases have operationalized Simon's (1957) intriguing, but ambiguous, notion of bounded rationality, in a manner that can be realized in experimental demonstrations.

Starting in the late 1980s, a newer group of biases became part of BDR. In its early years, BDR focused almost exclusively on cognitive errors, or errors that have their root in how we process information. Many of the newer biases proved not to fit into the framework of

computational limitations. Starting in the late 1980s, research emerged that suggested errors that we make as a result of motivational biases. These biases include 1) positive illusions - the tendency to view oneself and the world more positively than reality suggests (Taylor, 1989, Taylor and Brown, 1988), 2) egocentrism, or the self-serving ways in which people interpret fairness (Babcock, Loewenstein, Issacharoff, and Camerer, 1995; Babcock and Loewenstein, 1997; Diekmann, Samuels, Ross, and Bazerman, 1997), and 3) how we make biased decisions in order to avoid regret (Kahneman and Tversky, 1982; Kahneman and Miller, 1986; Medvec, Madey, and Gilovich, 1995). Note that these sorts of biases broaden the scope of inquiry to include self-concept, emotions, and social judgment. As a consequence, they draw BDR closer to mainstream cognitive and social psychology. Before bringing out the implications of this trend, we turn to some challenges to and criticisms of BDR.

### Challenges to Behavioral Decision Research

Given the successes of BDR one might imagine that it would be relatively immune to criticism. Not so. Criticisms of BDR have come in waves during its development. From Simon's work in the 1950s through the 1970s, BDR was discounted or ignored by economists - they continued to hold on to the rationality assumption (Lewin, 1996). Part of this reaction can easily be explained by self interest. However, part also is attributable to the ability of economists to explain BDR's arguments and evidence within their existing, parsimonious paradigm. Bounded rationality to an economist was simply rationality subject to search costs. Similarly, Tversky and Kahneman's (1974) heuristics - availability, representativeness, and anchoring and adjustment - were specific ways in which people rationally used heuristics to simplify search - the net savings in search costs, they argued, outweighed the benefits of more complete search. On this view BDR



could be relegated to a footnote qualifying the basic picture of optimal/rational behavior.

A breakthrough came as a result of Kahneman and Tversky's (1979) Prospect Theory, which showed (among other things) that choice could be dramatically affected by the frame in which information was presented. The frame was normatively irrelevant from an economic model. Prospect Theory was not easy to account for within the standard economic paradigm. Since 1980, there has been a dramatic increase in interaction between psychologists and economists over how to describe human decision processes. From a Kuhnian (1970) perspective, the economic model has been in crises, and the decision literature has received significant attention. This success was quickly followed by a new set of challenges to BDR.

Much of the attention received by BDR is a direct reflection of its ability to specifically critique our decisions as deviations from rationality (interestingly, a claim rarely made by Kahneman and/or Tversky). But this focus on departures from rationality was what made the literature so popular. In addition, others were quick to label BDR's biases as deviations from the norm of optimality (Bazerman, 1998). The label "deviation from rationality" focused the literature on mistakes that were being made. Yet, some scholars did not share this assessment of human behavior. Some continue to argue that biases provide net benefit, and still others have argued that these biases exist only in the minds of experimenters.

Taylor and Brown (1988) argue that biases enhance and protect self-esteem, increase personal contentment, help people persist at difficult tasks, and facilitate coping with aversive and uncontrollable events. Taylor (1989) even suggests that biases are beneficial to physical and mental health. Seligman (1991) proposes that salespeople be selected based on the magnitude of a specific bias - what he calls learned optimism. He argues that unrealistically high levels of optimism are useful for maintaining persistence in a sales force. We believe that while empirical

evidence exists for these specific positions, it may not be sustainable to continually fool oneself. This view is shared by a growing number of scholars who caution that biases are likely to have a negative impact on learning, on the quality of decision making, and on responses to organizational crises ("the oil in the water isn't really that big of a problem"). Biases can also contribute to conflict (impasse, divorce, and war) when decision-maker biases lead them to believe in the accuracy of their view (Brodt, 1990; Kramer, Newton, and Pommerenke, 1991; Kramer, 1995; Tyler and Hastie, 1991).

Another set of criticisms suggests that the BDR paradigm has overstated the magnitude of deviations from rationality, and even questions whether the deviations exist (Gigerenzer, 1991; Cosmides and Tooby, 1994; Gigerenzer and Hoffrage, 1995). Gigerenzer (1991) shows, for example, that while people may be overconfident in each specific judgment, they know about their overconfidence. Thus, while people may only be 60% accurate when make judgments for which they claim 95% confidence, they do not expect 95% of their 95% judgments to be accurate. As a result, Gigerenzer questions whether people are actually overconfident, or whether the specific methodologies of BDR make them appear so. From an evolutionary perspective, Cosmides and Tooby (1994) show dramatic reductions in biases when BDR problems are reframed in ways that make evolutionary "social sense". These researchers provide interesting data on factors that can reduce bias. However, our assessment is that these results sometimes are overstated to the extent that they are taken as contradicting the fundamental results of BDR (see Kahneman and Tversky, 1996, for a response to Gigerenzer and others). Rather, consistent with many BDR advocates (e.g., Dawes, 1997), we believe that these studies have simply shown that biases do not always exist or that various factors may work to reduce (or amplify) their magnitude. Not a single anomaly that has emerged in the BDR literature has been "destroyed" or eliminated by these

attacks (Camerer, 1995).

Experimental economists (e.g., Kagel and Levin, 1986) suggest that experience may eliminate or at least attenuate decision bias because performance feedback can correct the inappropriate use of information and decision heuristics. Some experimental economists see bias researchers as artificially creating effects that are limited to one trial demonstrations. Experimental economics is filled with demonstrations of performance improvement through multiple trials - with convergence on the economically rational solutions (Kagel and Roth, 1995). In contrast, we suggest that much of experimental economics consists of contrived experiments created in order to show convergence. In fact, many BDR results do live through shockingly high levels of feedback and multiple trials (Ball, Bazerman, and Carroll, 1991; Camerer, 1995). Furthermore, it is doubtful that real-world experience would provide decision makers with superior information or even useful feedback in many cases (see Einhorn and Hogarth, 1978, for a nice analysis of information and the severely limited feedback associated with many selection contexts). Learning from experience requires accurate and immediate feedback which is rarely available for a variety of reasons including:

(i) outcomes are commonly delayed and not easily attributable to a particular action; (ii) variability in the environment degrades the reliability of the feedback; (iii) there is often no information about what the outcome would have been if another decision had been made; and (iv) most importantly, decisions are unique and therefore provide little opportunity for learning...any claim that a particular error will be eliminated by experience must be supported by demonstrating that the conditions for effective learning are satisfied. (Tversky and Kahneman, 1986)

In fact, research has shown that most of the BDR effects tested occur with real people (that is, nonundergraduates), with multiple trials available for learning, and with rewards for successful performance. In virtually all cases, these biases are robust to the tests that critics have provided (Bazerman, 1998). In short, the heuristics and biases component of BDR appears to be on firm empirical ground. We believe that concerns which are more to the point focus on the need and desirability of greater integration of BDR with the rest of psychology and going beyond a negative definition of BDR (an issue to which we now turn).

### A Psycho-Central Framework

The previous section summarized past critiques of BDR. Generally we see BDR as very successful in responding to these challenges. We believe, however, that there is a more important set of criticisms coming from cognitive psychology, social cognition (Fiske and Taylor, 1991), and naturalistic decision research (Connolly and Koput, 1997; Messick, 1999). While the first two listed are common areas of psychology, the latter is an emerging area of organizational behavior that argues for a variety of new methodologies for studying decisions in organizations. Like cognitive psychology and social cognition, it critiques BDR for rarely describing decision *processes*. (Although some researchers in BDR have withdrawn from the backdrop of economics in favor of analyzing the component processes and subprocesses associated with decisions (Wallsten, 1980; Payne, Bettman, and Johnson, 1992; Busemeyer and Townsend, 1993), our present focus is on critiques and improvements to the heuristics and biases approach). It argues that the primary focus of BDR tells us what decisions makers don't do (act rationally), rather than describing what decision-makers actually do. This analysis may not ring true in that heuristics like anchoring and adjustment, availability and representativeness seem to be describing processes.

However, it is often the case that inferences about process are based on the outcomes of decisions with little direct evidence concerning the processes themselves. That is, the process analysis may not be especially deep.

This set of criticisms points out that the field of BDR has, in a sense, been defined by what it is not, by its departure from standard economic assumptions. This negative definition has come at the cost of (implicitly) agreeing with certain assumptions of the economic rationality framework. It is as if BDR were trying to defeat economic theorists at their own game--BDR has done so (in our opinion), but the rules of the game continue to be set by the economic framework. An associated consequence of this debate is that the processing side of decision making has received correspondingly less attention.

We embrace the criticism of the failure of the heuristics and biases tradition to describe how decisions are actually made. A better understanding of decision processes has the potential to inform BDR concerning how to eliminate biases and improve decisions. In addition, there are insights from allied areas that are beginning to inform BDR research, such as interactions between social cognition and BDR (Nisbett and Ross, 1980; Morris and Larrick, 1995; Bazerman, Curhan, and Moore, 2000). These researchers have begun to expand the reach of BDR by adding social psychological variables, including emotions, fairness, relationships, and the broader social context (Thompson, 1998). More important for our purposes, this research holds the promise of creating a better connection between other areas of psychology and behavior decision research. The goal is to make these connections while maintaining the success of the heuristics and biases approach in creating usable knowledge that will provide insight to a variety of applications and lead to useful prescriptions.

On our analysis, current trends in decision making are beginning to question certain core

assumptions that serve to define both what problems are worthy of study and how they should be studied. In the next section we overview behavioral decision research in terms of the level of analysis that it typically uses, and suggest that a new, broader definition of behavioral decision research can develop by broadening the levels of analysis explored.

### Levels of Analysis

One way to organize the potential of BDR is to think in terms of levels of analysis. Palmer and Kimchi (1986) suggest that information processing theories vary in their abstractness and can be understood in terms of what they call “recursive decomposition.”<sup>1</sup> On this view, any complex event at one level of description can be specified more fully at a lower level of description by decomposing it into a number of components and processes that specify the relationships among these components. Research in BDR can also be seen as involving different levels of analysis (See also Busemeyer, Hastie, and Medin, 1995 for a discussion of levels of decision making). Note that recursive decomposition does not dignify any one level of analysis as special or qualitatively different from other levels. For this review we would like to make a distinction and, for this purpose draw on the analysis of levels made by Marr (1970).

Although David Marr’s research was in computational vision, he argued that the functioning of any intelligent system is best understood in terms of three distinct levels of analysis. These levels differ in their abstractness and in the sorts of questions they attempt to address. The first and most abstract level is known as the Computational Level.<sup>2</sup> At this level the central question is what is the system trying to do. In other words what are the goals that will dictate which sorts of information are relevant and which actions appropriate? For example, in the case of a system like a thermostat the goal is to maintain the temperature within a certain range. This goal

determines what information is relevant. For example, the prevailing temperature is very likely to be relevant and whether the Chicago Bulls have a televised basketball game in the evening is very likely to be irrelevant. There is not necessarily a single goal; for example, the goals may include minimizing the variance of the temperature or perhaps the time spent processing information relevant to the goal. The latter consideration begins to touch on the second level of analysis which Marr called the Algorithmic Level.

The Algorithmic Level is concerned with what computations are performed and the theory that determines or describes how these computations achieve the goals of the system. In the case of a thermostat this might be as simple as determining the current temperature, the desired temperature, and the (signed) difference between them, linked to some procedure for reducing the difference. Note that this level of description is still quite abstract and would allow a wide range of instantiations. For example, temperature could be measured by a variety of mechanisms and the output could be in either Fahrenheit or Centigrade. In the case of decision making the Algorithmic level might be concerned with how utility is determined and how the computational and decision processes function to maximize utility.

The question of how the overall goals and associated computations are actually performed by a physical device addresses Marr's Implementation Level. In the case of the thermostat, the implementation level refers to the mechanical device that performs the function; for much of psychology, the implementation level is the nervous system (though to be true to the complexities involved in decision making, a full account would need to include other people – [e.g., advisors], as well as decision aides like computers that may perform relevant computations).

We do not see BDR research as neatly organized into these three levels. Important distinctions between different levels of processing and analysis may be lost within Marr's broad

Algorithmic Level and there may not be any BDR that conforms literally to Marr's

Implementation Level. Nonetheless, we think that Marr's Computational Level questions are important for BDR and that it is also important to distinguish among levels of processing that may range from fairly abstract algorithmic descriptions to those that quite specific and begin to approach the implementation side of things. Therefore, we review well borrow Marr's Computational Level but then substitute Palmer and Kimchi's recursive decomposition for Marr's Algorithmic and Implementation levels. We first turn to a Computational Level analysis of BDR and follow it with a review of newer work that focuses on probing more deeply into decision making processes, including questions about how decisions get implemented (note we are not using "implemented" in Marr's literal sense but rather in a relative sense).

At the Computational Level, rational economic theory dictates that what people are trying to do is to maximize (subjective) utility. The theory does not try to describe what people will value; it simply says that different options will have different value or utilities and that the decision maker will choose the option with greatest utility. At first impression the theory seems circular; it says we choose A over B because A has more utility than B. But how do we know that A has more utility than B other than by the choice itself?

But the theory does have content. First of all, it implies that different measures of preference should agree with each other, If I prefer A over B, I should be willing to pay more money for A than B and require more money to give up A than B. In addition, this assumption requires that indirect measures of preference be consistent---if I prefer A to B and prefer B to C then the theory says I will necessarily prefer A to C. (In contrast, research on decision making has identified a variety of conditions under which people systematically violate both of these measures of consistency [e.g., Dawes, 1997]).



Economic theory is relatively silent on with respect to Marr's Algorithmic Level and this is as it should be, in our opinion. Attempts to break that silence appear to lead to error and prejudgment. Consider what happens when the Computational Level is used to make inferences about processing as in the following example from Frisch and Clemen, 1994:

“..Imagine that a person has a \$1000 balance on his or her VISA card and pays 17% annual interest. Imagine that this person also has \$2000 in a savings account earning 4% interest. From the perspective of utility theory one would conclude that the utility to this person from having money in a savings account (e.g. feeling of security) outweighed the cost of paying the high interest on the VISA bill.”

As Frisch and Clemen note, however, whether this choice reflects the results of an explicit assessment of value or utility should an empirical question, not a given. They suggest that it may reflect a habit of keeping the money in savings and it is easy to think of other possibilities. For example, 1) the person in the example may not have remembered that they had money in a savings account during times when they paid their bill, 2) the savings money may have been a gift from a relative and may be linked to a moral prohibition from using it to pay current expense bills or 3) the credit card bill may have been produced by some impulsive purchases and the person in question may have decided to punish himself or herself with the high interest payment (or they may be protecting themselves from future impulse buying by leaving a balance near their credit limit). The point is that, when treated as a process model, utility theory prejudges the basis for decisions and forecloses on the exploration of other kinds of motivations. Post hoc, one can almost always make up utilities to explain examples (e.g. see Machina, 1982, for an interpretation

of framing effects in terms of utility), but that robs the theory of explanatory power. We believe that it is better to treat utility theory as an abstract Computational Level theory.

The distinction between the Computational and Algorithmic levels is also roughly analogous to the distinction in BDR between prescriptive and descriptive accounts of decision making. Prescriptive analyses specify what one should do given certain goals while descriptive accounts focus on what people actually do which often falls short of the prescriptive ideal. Utility theory has had and will continue to have important applications as a normative model in particular contexts where the relevant variables have been isolated and assumptions critical to its use are satisfied. But it is not a process model.

Summary. One reading of the field to date is that economic rationality provides the Computational Level analysis with BDR filling in the Algorithmic Level with models of how people are not optimal. But many people in BDR do not see themselves as showing that human beings are (so to speak) just a bit less than the angels---they see heuristics and biases as more serious and less easy to attribute solely to computational limitations. And some of them worry about the field of decision making being only loosely tied to the rest of psychology. Perhaps the accumulation of observations and analyses is beginning to raise the question of whether utility theory is being asked to do too much work and whether a different Computational Level analysis would provide better guidance for the field of BDR. For example, utility theory is silent with respect to the processes associated with the formulation of a decision maker's goals that precede decision making. People are also asking whether all the various goals and considerations associated with decisions can be mapped on to a common utility function, or are there goals that are not additive or substitutable? Are there different *kinds* of decisions? If decision making is simply computations of utility and selecting the option with the highest utility, then why do people

sometimes experience intense conflict over a decision?

There are also clear signs that BDR is starting to attend to less abstract levels of analysis as a source of ideas about processing mechanisms. For example, there are some intriguing observations on the effects of frontal lobe damage on decision making (e.g., Damasio, 1994) and a neural network model or two representing decision making processes at levels that are at least inspired by the nervous system (Usher and Zakay, 1993; Leven and Levine, 1996; Thagard, 1989; Holyoak and Simon, 1999). In addition, Loewenstein (1996) and others recently have begun to reintroduce emotional and visceral factors as key variables that affect decision making (social psychologists would point to Kurt Lewin (1935), to Neal Miller (1944) and perhaps even to Freud for historical precedents). Researchers are also taking more seriously the fact that people sometimes talk about decisions in terms of gut feelings or as their head telling them one thing while their heart tells them another. For example, work is beginning to appear on the role of thoughts versus feelings in decision making (e.g., Wilson and Schooler, 1991; Loewenstein, 1996). In short, there is something of a trend to supplement the view of decision makers as computational devices with the perspective that humans beings are biological and social organisms whose decision processes may be shaped by biological and social processes.

Our diagnosis of the current situation is as follows. Rational choice theory and BDR have gone somewhat hand in hand, with BDR gaining stature and significance for its critiques of economic rationality as a descriptive model. Although rationality has served as a good foil, as a Computational Level theory it is too abstract to guide the detailed development of process models of decision making. As the field of BDR increasingly shifts to a concern with process it should look to (and develop) less abstract computational level frameworks for guidance.

In the rest of this paper, we outline the potential of a productive interaction between

cognitive and social psychology and BDR. We focus on the implications of studies for Computational Level and other levels of analyses (ranging from problem definition to implementation aspects of decisions) for better understanding decision making. We will not presume to lay out new theories for decision making; rather, we will describe some of the strong currents that are motivating a rethinking of the Computational Level and probing more deeply into processing mechanisms. Although a consensus has yet to emerge with respect to a new framework theory, there are a number of common elements that run through the research themes that we will describe.

### Toward a New Computational Level Framework of Decision Making

Goals. Our colleague, David Messick (1999), recently came forward with his embarrassing tale of attending a multi-family retreat, which featured a ping-pong tournament. It was not until David was too far ahead in the finals against a 14 year-old girl, that he realized that the goal of the tournament (from the adults' perspective) was for the parents to lose early, allowing children to end up in the finals. Dave won the tournament, but overall was very ineffective - by his own definition. As he now explains, (Messick, 1999), the problem was not with determining when to spin the ball or with any aspect involving ping-pong skill. He simply had a fundamental misunderstanding of the meaning and goals of the tournament. Messick uses this story as a springboard to discuss systematic ways in which decision-makers misconstrue the goals of their decision processes. He goes on to analyze how the order of information, labels, causal explanations, metaphors (Neustadt and May, 1986), timing (Morris, Sim, and Girotto, 1998; Shafir and Tversky, 1992), and social punctuation (Kramer, 1995; Tversky and Kahnmean, 1986) can all affect the definition of the problem to be solved.

In the language of economics, Messick failed to grasp what would provide utility. As noted earlier, economic theory is consistent with a very abstract level of analysis that generalizes over content and concerns itself solely with likelihood and desirability or utility (Kahneman, 1991). Utility acts as a common currency that allows one to integrate over the various positive or negative facets of alternatives. Making a decision then becomes a simple matter of determining the option that maximizes utility. As we have been suggesting, economic models of utility maximization are increasingly seen as unsatisfactory because their level of analysis is too abstract. A richer Computational Level analysis of decision making is needed. We now turn to evidence that argues against the simple additive nature of goal integration assumed by the concept of utility. Instead, the nature of making a decision often creates meaning that becomes part of the goal set of the decision-maker.

Multiple Goals - Limitations of a Common Currency Framework. Thaler's (1985) work on mental accounting supports a level of analysis in terms of multiple goals rather than a single, overarching goal of maximizing utility. If so, then an important Computational Level function is to satisfy and coordinate these goals. The distinct categories associated with mental accounting may be linked to distinct goals (Henderson and Peterson, 1992). Brendl, Markman and Higgins (1998) have developed this perspective in some detail. They argue that currently active goals set up mental accounts and that gains and losses are weighted into these accounts proportionally to their representativeness with respect to the goal. As one test of the idea that currently active goals dominate accounting they presented the following scenario: "Imagine two students are visiting a gambling casino. In front of the casino Student A finds \$25 cash and puts the money in his wallet. Student A and Student B pay the \$25 entrance fee and enter the casino. Inside the casino Student

B finds \$25 cash and puts the money into his wallet. Both students do not know yet whether to gamble or not."

Participants were then asked which of the two students would be more likely to place a \$25 bet that had a 50% chance of winning \$25 and a 50% chance of losing \$25. Consistent with our intuition, Brendl et al. predicted that Student B will be selected by participants as more likely to gamble. Their logic is that when Student A pays the \$25 entry fee he is likely to assign the \$25 he found to that account and goal. Student B, in contrast, has already satisfied the goal of gaining admission at the point he finds his \$25 and that money is "free" to be placed into the gambling account. In agreement with this analysis 72% of the participants judged that Student B was more likely to take the bet.

Mental accounting categories may have the function of preventing a focus on one goal at the expense of other goals. As Henderson and Peterson (1992) note, goals, like taxonomic categories, appear to be hierarchically organized and, therefore, principles associated with categorization research may apply to mental accounting and decision making. Theory and data in the area of categorization suggest that inter-property relationships often provide conceptual coherence (e.g., Murphy and Medin, 1985). The same may hold for decision making. For example, if a cost is seen as causally related to a benefit, then the cost and benefit are more likely to be integrated into the same mental account (Hirst, Joyce, and Schadewald, 1994).

There are several other attacks that have been made on the common currency component of utility theory. One focuses on the decision making process itself. A generalization from much of the BDR literature is that people prefer making comparisons at a more concrete level than implied by utility theory (Payne, Bettman, & Johnson, 1992). That is, they often prefer to analyze

options, component by component rather than by independent overall assessments. People also tend to give more weight to directly comparable differences. For example, Slovic and MacPhillamy (1974) asked participants to rate students who have hard scores on one common dimension (e.g., English skills) and one unique dimensions (e.g., quantitative skills for one student and achievement need for the other). When a dimension was common, it received much more weight than when it was unique. Instructions not to overweight the common dimension did not reduce this effect.

In a related study Tversky and Kahneman (1986) gave people choices between bets involving boxes of marbles that varied in color and the outcome associated with them (people were told that they would sample one marble from the box they chose and receive its associated outcome). Tversky and Kahneman expected that people would align corresponding colors to evaluate the overall attractiveness of the boxes (see Table 1). They varied the payoff and color combinations such that the identical overall structures would lead to different patterns of comparisons in different groups. Assuming that corresponding colors are compared it is easy to see that Option A dominates Option B but the realignment associated with Options C and D makes this much less obvious. In fact, if people focus on the biggest difference in outcomes (associated with drawing a green marble), they might choose D over B. And the data were in accord with these expectations---participants given a choice between A and B picked A 100% of the time; in contrast, participants given this same distribution of gains and losses picked D over C 58% of the time. In short, comparability of components has robust effects on choices, a finding that is inconsistent with people simply adding up pluses and minuses to determine overall value (see also, Tversky and Sattath, 1979, Markman and Medin, 1995, and Shafir, 1995 for related evidence).

---

Insert Table 1 about here

---

There is also evidence that judgment is not always a monotonic function of the value or utility of each component considered separately. Anderson and Alexander (1971) found that adding moderately favorable information to highly favorable information produced less favorable responses. Even more striking are the observations of Kahneman and his associates (Kahneman, Fredrickson, Schreiber, and Redelmeier, 1993) in a task where participants underwent the painful experience of holding their hand in iced water for vary periods of time. The coldness of the water was varied between 1 (moderately painful) and 10 (extremely painful. Participants rated the experience as more unpleasant when they received the coldness levels [2,5,8] than when they experienced the levels [2,5,8,4]. That is, adding a moderately painful sample actually decreased unpleasantness judgments. These results are problematic for any model that ascribes a single valence to a piece of information since the same, single facet can either increase or decrease an overall impression (a “4” by itself was rated as unpleasant).

One could attempt to account for these results by assuming that utility is based on an average of components rather than a sum. Aside from the fact that this move would not handle the details of the Kahneman et al. findings, such a move is unmotivated by utility theory and one would need a meta-theory describing which situations trigger various kinds of integration



functions. Note that, in principle, judgments and experienced utility could have been a simple additive (or at least monotonically increasing) function of the components, in which case the Computational Level of analysis associated with utility theory would be a good guide to predicting behavior in situations like those just considered. But it didn't turn out that way. Utility theory can be saved by the sort of second order analyses that we have frequently seen, but a theory that can predict every conceivable pattern explains nothing. Hence our claim that the theory just is not very useful for much of the current work in decision making.

Even when one has a prototypical homogeneous entity (i.e., money) people do not appear to treat it as homogeneous. Consider the following scenario taken from Thaler (1985, p. 199). “Mr. and Mrs. J have saved \$15,000 toward their dream vacation home. They hope to buy the home in five years. The money earns 10% in a money market account. They just bought a new car for \$11,000 which they financed with a three-year car loan at 15%.” This example shows a clear violation of the fungibility of money. Normatively speaking, Mr. and Mrs. J should use money from their savings account to pay for the car. In contrast, people often create budgets which impose constraints on the transfer across kinds of accounts. Thaler argues that budgeting in this case is a response to self-control problems and that if the vacation home account is drawn down, it might not be repaid (whereas the bank will insure that the car loan is repaid). In a similar vein, people typically do not spend money they have received as a gift the same way they spend money they earn. Rather it tends to be “segregated” and earmarked for something one would not ordinarily purchase (see Thaler, 1985; Thaler and Johnson, 1990 for other examples and empirical demonstrations). These examples violate standard economic theory which treats money just as money.

In brief, a variety of observations go against the notion that utility is a homogeneous entity

that acts as a common currency. To be clear about our thesis, we are not claiming that utility theory cannot represent utilities in such a way as to describe people's decisions, after the fact. Nor are we claiming that competing demands or goals somehow escape the need to be integrated in the service of action (though we expect that the integration is almost always more complex than summing constituent utilities and selecting the option with the highest total). Instead, we are saying that, as a Computational Level theory, utility models provide no rationale for considering multiple goals and, therefore, offer little insight into decision processes. To be sure, multiple facets each with their own utility may go into overall utility but nothing in utility theory distinguishes between integrating facets within versus across goals or domains. In short, arguing that people act to maximize their utility is a claim too abstract for the purpose of understanding how decision makers coordinate competing goals.

Decision making and meaning. Consider the following ultimatum game (Guth, Schmittberger, and Schwarze, 1982; Roth, 1991). Participants are paired off (call the player A and B) where one person (A) is allowed to propose an allocations of some resource and the other (B) must accept it or reject it. If the offer is accepted both players receive their allocation; if B rejects the offer, neither player receives anything. For example, the resource may be \$10 and A may propose to take \$9 and give B \$1. From one point of view the B's decision is quite simple--B should accept any nonnegative outcome because it is better than nothing.

To the contrary, B's typically reject such an unequal distribution on grounds that it is not s is a 50-50 split, even though the participants do not know each other and are unlikely to interact again in the future. These basic ultimatum results have been replicated easily over time (Bazerman, 1998). These results also question the notion of decision-makers having clear, a priori set of concerns. Rather, it appears as if the act of making the

decision creates meaning to the outcomes. Thus, the rejection of a profitable offer allows the rejecter to place value on harming the party making an unjust ultimatum offer. Thaler (1985) suggested that transactions involve two kinds of utility, acquisition utility and transaction utility. The former depends on the value of the good in itself and the latter refers to the perceived merits of the deal. For the above example, the acquisition utility is positive but the transactional utility may be negative. Interestingly, notions of fairness, trust, and reciprocity which have no role in standard rational analyses of social dilemmas appear to allow groups to achieve levels of outcomes that are “better than rational” in that exceed what is predicted by rational choice analysis (Ostrum, 1998).

More generally, decisions may have the property of conveying information to both the decision maker and to others. Medin, Schwartz, Blok and Birnbaum (1999) argue that potential meanings are taken into account by the decision-maker and ,as a consequence may affect decisions themselves. For example, undergraduates asked to imagine that they were working in a restaurant say that they would rather receive no tip than a tip of one cent. They considered a penny tip as an insult and noted that the absence of a tip might reflect forgetfulness or lack of money on the part of the patron rather than dissatisfaction with the service. In short, a penny tip is not just money--it’s a message.

Medin et al. studied people’s ideas about the exchangeability of entities that in some cases might have symbolic (e.g., a wedding ring) and/or sentimental value. They found that people were very reluctant to exchange such things for objectively comparable entities (an identical wedding ring) even when offered a large monetary incentive to do so. A typical justification was “It’s a meaning issue, not a money issue.” In other conditions participants made judgments about fictional people making hypothetical exchanges. For example, participants rated a person who

gave away his dog to help cheer up patients at a children's hospital quite favorably. But when this scenario also involved the person receiving a thousand dollars in exchange, the person was rated very unfavorably. Justifications suggest that the money received called into question why the owner had given up his dog and participants found the idea of giving up one's dog for money repugnant. In short, judgments about exchangeability were strongly constrained by the meaning associated with the decision.

A focus on meaning is also consistent with the idea that there may be kinds of decisions with different principles and strategies associated with a given kind. These sorts of influences are most readily seen in studies where (more or less) the same context is involved but cues are present that bias the interpretation of the context in one way or another. For example, Larrick and Blount (1997) presented participants with game that could be seen either as a social dilemma or as involving fairness (an ultimatum game). They varied whether the situation was described in terms of "claiming" versus "accepting or rejecting" offers. Larrick and Blount reasoned that the accept or reject framing is more compatible with an ultimatum game than with a social dilemma. They found that the description affected both the offers made and their likelihood of being accepted---the framing in terms of accepting versus rejecting was associated with less willingness to accept small shares than the framing in terms of claiming. In the same vein Tenbrunsel and Messick (1999) presented participants with a commons game involving businesses that could cooperate to control pollution. They found that introducing small penalties for polluting actually produced less cooperation than a control (no penalties) condition. Tenbrunsel and Messick also collected measures of how people interpreted the decision task. Without penalties, most people thought of the game as requiring a personal or ethical decision. When penalties were present, in contrast, most people saw the game as involving a business decision. Decision behavior was more a

function of how the context was interpreted than of objective costs and benefits.

In perhaps the most elaborate theoretical statement concerning meaning effects, Fiske and Tetlock (1997) proposed that the rules governing decisions and exchanges are a function of the type of social relationship involved. On their analysis there are four types of social relationships: communal sharing, authority ranking, equality matching, and market pricing. Communal sharing is categorical in character--members do not “keep track” of favors nor is direct reciprocity involved. For example, a father who helps his daughter move into a new apartment would not expect to be compensated (monetarily or otherwise). Friendships typically correspond to equality matching where there is something of a loose reciprocity and problems may arise if exchanges get too far out of balance. For example, there is nothing wrong with going to one’s grandparents house for Sunday dinner week after week but it would be very unusual and perhaps strange to go to your friend’s house week after week for dinner (of course, one can imagine conditions under which this might happen and be seen as appropriate, but the norm is that dinner invitations are reciprocated). Finally, market pricing involves a precise accounting associated with exchanges, the prototypical example being goods and services being exchanged for money.

As Fiske and Tetlock point out, trying a form of exchange that does not map on to the type of relationship will result in misunderstandings or be treated as insulting. It will not do to respond to a nice dinner in a restaurant by saying “Thanks, you must come to our house for dinner soon” or to respond to an equally nice dinner at a friend’s house by saying “Do you take credit cards?” Nor is the problem simply one of introducing money---it would be also be a serious social gaff to offer to mow your friend’s lawn as a form of immediate reciprocation for the dinner. Note that, within the Fiske and Tetlock framework, utility and substitutability are not domain wide but rather are a function of the social context. Money or bartering may not buy you much if other

people are unwilling to enter into a market pricing relationship.

People often employ specific strategies to overcome or circumvent attributions that might otherwise be made. For example, one of our colleagues works at home on Wednesdays and has a policy of never coming in to the office on Wednesdays. This policy serves several functions, one of which is to avoid the appearance of explicitly weighing time working at home on a specific occasion against some concrete alternative like attending an “important” meeting. Instead, we interpret her behavior as a strategy for achieving some long-term goals by deliberately avoiding a calculation of costs and benefits each Wednesday.

Implications. The above research suggests that decision making has both a meaning component and a social component. Decision making may involve multiple goals, including conveying information to others (and oneself). Meaning is often about goals and goals and meaning may interact to determine decisions. We read generalizations such as these as indicating that utility theory is ineffective as a Computational Level framework for the area of decision making because it fails to make contact with or serve as a positive framework for much of contemporary research.

The final section of our review turns to research that focuses on working out the details of more realistic process models of decision making. Here we will see that the idea that the idea that people have multiple goals is closely paralleled by work suggesting that the decision maker may have competing motivations. Within the framework of utility models, choices are said to reveal preferences and it wouldn't be sensible to suggest that sometimes choices are inconsistent with preferences. Some of the research to be considered next which focuses on the internal conflict associated with decisions calls into question the idea that choices necessarily reflect preferences (see Lazar, 1999 and Lerner & Tetlock, 1999, for two distinct perspectives on this

issue).

### The Implementation Side of Decision Making

Starting from the point where the individual has determined the best decision, implementation would seem relatively straightforward. In contrast, ample evidence now argues that details of the processes associated with decision making need to be taken very seriously. To make that point, many researchers (e.g., Thaler and Sherfin, 1991; Schelling, 1984) have retold the story of Ulysses, where Ulysses knows that he will soon encounter the Sirens. The Sirens are mythical female "enchanters" who lure seafaring men to their subsequent death by singing to them sweetly from their flowering meadow. Ulysses orders his men to put wax in their ears to block out the tempting voices of the Sirens. For Ulysses himself, he orders his men to bind him with ropes to the ship's mast, and to order them in advance not to release him, no matter how much he begs, until after they pass safely by the Sirens. This way, Ulysses will have the pleasure of listening to the Sirens without being trapped by them.

Given the desire and decision to live, why are all of these precautions needed? Ulysses' problem is an example of a group of problems that affect our decisions on a regular basis. These conflicts are called the multiple selves problem (Sidgwick, 1874; Strotz, 1956; Ainslie, 1975, 1992; Sen, 1977, Winston, 1980; Schelling, 1984; Elster, 1985; Shefrin and Thaler, 1988; Walsh, 1994, 1996; Loewenstein, 1996). Recently, Bazerman, Tenbrunsel, and Wade-Benzoni (1998) have labeled these two selves the *want self* and the *should self*. In the language of the current paper, the should self represents the decision coming out of the algorithmic stage of the decision process, while the want self represents transient concerns that pull the decision maker away from the decision created at the algorithmic stage. Bazerman et al.'s (1998) distinction of want/should

is broadly compatible with Loewenstein's (1996) who argues that much human action is based on visceral or transient responses that go against long-term self-interest. Both Loewenstein (1996) and Bazerman et al. (1998) discuss the advantages and disadvantages of each of the two selves. Bazerman et al. (1998) then move on to offering prescriptions about how to use input from each of the selves to formulate a wise strategy.

Psychologists have long studied related internal battles. Higgins, Tykocinski and Vookles (1990; Higgins, Roney, Crowe and Hymes, 1994; Higgins, 1987; for a more formal model see Busemeyer and Townsend, 1993) specified a number of different facets of the self that can create internal conflict, including actual, ideal, ought, can, and future selves. Higgins (1987) distinguished between two "actual" selves -- the kind of person an individual believes she actually is and the kind of person an individual believes that others think she is. Erikson (1950; 1963), Lecky (1961), Mead (1934), and (Wylie, 1979) describe a variety of other dichotomy of the self (Markus & Nurius, 1987). In addition, James (1890; 1948) distinguishes the *spiritual* self and the *social* self. Rogers (1961) distinguished between what a person should be according to a normative standard and a person's preference about what she would ideally like to be. Schafer (1967) and Piers and Singer (1971) elaborated on Freud's (1923/1961) distinction between the ego and superego. Similarly, research on deliberate versus automatic processing and mental contamination (Wilson and Brekke, 1994) has distinguished between unwanted judgments, actions, and moods that deviate from the preferred state.

Bazerman et al. (1998) further argue that it is possible to predict a priori when each of their two selves, want and should, will dominate. Specifically, they summarize evidence that shows in risk-free decisions, the number of options presented to a decision maker differentially triggers one of the two "selves." "Want" dominates in situations where one concrete option is



under consideration. In contrast, the more reasoned "should" dominates when individuals confront multiple options at the same time. Essentially, Bazerman et al. (1998) claim that we have a more difficult time implementing the preferred algorithmic solution when we lack concrete options for comparison, and suggest that this results from the power of transient concerns in such environments.

In a provocative book Frank (1988) argues that folk theories about “passions” support an important and useful role for emotions in decision making contexts, precisely because emotions do not necessarily embody a rational calculation of self-interest. Frank gives the example of a person who is upset because the person next door lets his dog wander into neighboring backyards and doesn’t clean up any “messes” his dogs leaves. The angry person threatens to take the neighbor to court unless he controls his dog. The offending neighbor could assume that a rational person would analyze the costs and benefits of a court action and quickly realize that it would involve too much time and trouble to pursue a legal remedy. And therefore he could assume that he could safely continue to let his dog wander without needing to clean up after him. But, Frank reminds us, the person is very angry and the neighbor would need to take into account the fact that an angry person might ignore personal costs to punish the offender. Therefore, the neighbor might well decide that he had better keep his dog under control. In short, the person’s passion produces a desirable consequence, an outcome better than would be expected from cold calculations. In the same vein socio-biologists have suggested that passion, because it is hard to fake, provides an unambiguous signal for evaluating the status of romantic relationships.

Finally, recall that Damasio (1994) finds that people with frontal lobe damage show seriously impaired decision making despite having normal IQ. Space does not permit our going into his observations in detail but it appears that the damage is associated with diminished affect

and a corresponding inability to anticipate the negative aspects of risky decisions. This syndrome is as fascinating as it is sad. It may mark the beginning of a critical exploration of the brain and decision making.

Summary. This section is complementary to the previous one. At the Computational Level we suggested that it was more useful to think in terms of people satisfying multiple goals and making distinct *kinds* of decisions rather than aiming to account for decisions with the single goal of maximizing a homogeneous entity, subjective utility. From an implementational perspective it appears that there are multiple mechanisms or components serving somewhat different functions in decision making.

### Conclusions

One could imagine that this paper might be read as an attack on behavioral decision research for its narrow focus on the algorithmic level of processing. But that would be wrong. We think that BDR has been a remarkable success story, and perhaps all the more remarkable given its relative neglect of both the implementational side of decision making and a the need for a less abstract Computational Level analyses.

Our diagnosis is that utility theory has entered a state of encountering a set of clear threats within a Kuhnian (1970) revolution. Much as Kahneman and Tversky put the rationality assumption on the scientific table, we see research that questions utility theory's assumption of comparable and combinable goals as a descriptive model of behavior. We are not saying that utility theory is of no use; it has been and will continue to be an invaluable tool in a number of applied contexts. In addition, it remains the dominant model of goal specification until psychologists can more clearly provide a parsimonious theory of goals and their relation to kinds

of decisions. The burgeoning research on multiple selves, multiple goals, social relationships and meaning needs a framework that defines decision making by more than what it is not. We have tried to point to the ingredients for a more positive computational level theory, one that provides a better match for and better serves to organize research at more specific processing levels.

We have also tried to highlight and organize recent research which focuses on the question why decision-makers may not act on what they decide is the best possible decision -- the question of decision implementation. This is clearly an area where behavioral decision theory is well served through better connections to other areas of psychology, including physiological, cognitive, and clinical psychology. Quite simply, people often fail to do what they believe that they should do. A useful descriptive model of decision making should be able to provide insight into this inconsistency.

Behavioral decision research has been extremely effective from both practical and theoretical perspectives. But we think that a deeper understanding of the decision making process will increase the entry points for thinking about how to make decision makers more effective. We see a science in which behavioral decision research can help decision-makers better understand their goals, their decision process, and the implementation of these decisions, rather than simply the mistakes that are made at the point of decision. We hope that our review and analysis provides seeds for this new direction.

## References

- Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. Psychological Bulletin, 82, 463-509.
- Ainslie, G. (1992). Picoeconomics: The interaction of successive motivational states within the individual. New York: Cambridge University Press.
- Anderson, J. R. (1991). Is human cognition adaptive? Behavioral and Brain Sciences, 14, 471-517.
- Anderson, N. H., & Alexander, G. R. (1971). Choice test of the averaging hypothesis for information integration. Cognitive Psychology, 2, 313-324
- Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. Journal of Economic Perspectives.
- Babcock, L., Loewenstein, G., Issacharoff, S., & Camerer, C. (1995). Biased judgments of fairness in bargaining. American Economic Review, 85(5), 1337-1343.
- Ball, S. B., Bazerman, M. H., Carroll, J. S. (1991). An evaluation of learning in the bilateral winner's curse. Organizational Behavior and Human Decision Processes, 48, 1-22.
- Bazerman, M. H. (1998). Judgment in managerial decision making (4th ed.). New York: John Wiley & Sons.
- Bazerman, M. H., Curhan, J., & Moore, D. (2000). The death and rebirth of the social psychology of negotiations. Blackwell Handbook of Social Psychology.
- Bazerman, M. H., & Neale, M. A. (1992). Negotiating rationally. New York: Free Press.
- Bazerman, M. H., Tenbrunsel, A. E. & Wade-Benzoni, K. A. (1998). Negotiating with Yourself and Losing: Understanding and Managing Conflicting Internal Preferences. Academy of Management Review, 23, 225-241.

Brendl, C. M., Markman, A. B., & Higgins, E. T. (1998). Mentale Kontoführung als Selbstregulierung: Repräsentativität für zielgeleitete Kategorien [Mental accounting as self-regulation: Representativeness to goal-derived categories.] Zeitschrift fuer Sozialpsychologie, 29, 89-104.

Brodt, S. E. (1990). Cognitive illusions and personal management decisions. In C. L. Cooper & I. T. Robertson (Eds.), International Review of Industrial and Organizational Psychology. Vol. 5 (pp. 229-279). New York: Wiley.

Busemeyer, J. R., Hastie, R. & Medin, D.L. (1995). Preface. In J. R. Busemeyer, R. Hastie, & D.L. Medin, (Eds.), The Psychology of Learning and Motivation, Vol. 32. San Diego: Academic Press.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. Psychological Review , 100, 432-459.

Camerer, C. F. (1995). Individual Decision making. In J. H. Kagel & A. E. Roth (Eds.), The Handbook of Experimental Economics (pp. 587-703). Princeton: Princeton University Press.

Connolly, T., & Koput, K. (1997). Naturalistic decision making and the new organizational context. In Z. Shapira (Ed.), Organizational Decision Making. Cambridge, UK: Cambridge University Press.

Cosmides, L., & Tooby, J. (1994). Better than rational: Evolutionary psychology and the invisible hand. American Economic Review, 84, 327-332.

Dawes, R. M. (1988). Rational choice in an uncertain world. New York: Harcourt, Brace and Jovanovich.

Dawes, R. M. (1997). Behavioral decision making and judgement. In D. T. Gilbert & S. T. Fiske (Eds.), The handbook of social psychology, Vol. 2 (4th ed.) (pp. 497-548). Boston:

McGraw-Hill.

Damasio, A. R. (1994). Descartes' error: Emotion, reason and the human brain. New York: Putnam.

Diekmann, K. A., Samuels, S. M., Ross, L., & Bazerman, M. H. (1997). Self-Interest and Fairness in Problems of Resource Allocation. Journal of Personality and Social Psychology, 72, 1061-1074.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence illusion of validity. Psychological Review, 85, 395-416.

Elster, J. (1985). Weakness of will and the free-rider problem. Economics and Philosophy, 1, 231-265.

Erikson, E. H. (1963). Childhood and society (2<sup>nd</sup> ed.). New York: Norton. (Original work published 1950).

Fischhoff, B. (in press). Judgment heuristics. In R. Wilson & F. Keil (Eds.), The MIT encyclopedia of the cognitive sciences. Cambridge, MA: MIT Press.

Fiske, S. T., & Taylor, S. E. (1991). Social cognition (2nd ed.). New York: McGraw-Hill.

Fiske, A. P., & Tetlock, P. E. (1997). Taboo trade-offs: Reactions to transactions that transgress the spheres of justice. Political Psychology, 18, 255-297.

Frank, R. H. (1988). Passions within reason: The strategic role of the emotions. New York: Norton.

Freud, S. (1961). The ego and the id. In J. Strachey (Ed. and Trans.), The standard edition of the complete psychological works of Sigmund Freud, 19, 3-66. London: Hogarth Press. (Original work published 1923).

Frisch, D., & Clemen, R. T. (1994). Beyond expected utility: Rethinking behavioral

decision research. Psychological Bulletin, 116, 46-54.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. Psychological Review, 103, 650-664.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. Psychological Review, 102, 684-704

Goldstein, W. M., & Weber, E. U. (1995). Content and discontent: Indications and implications of domain specificity in preferential decision making. In J. Busemayer, R. Hastie, & D. L. Medin (Eds.), The psychology of learning and motivation, Vol. 32. San Diego: Academic Press.

Guth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. Journal of Economic Behavior and Organization, 3, 367-388.

Henderson, P. W., & Peterson, R. A. (1992). Mental accounting and categorization. Organizational Behavior and Human Decision Processes, 51, 92-117.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. Psychological Review, 94(3), 319-340.

Higgins, E. T., Roney, C. J. R., Crowe, E., & Hymes, C. (1994). Ideal versus ought predilections for approach and avoidance distinct self-regulatory systems. Journal of Personality & Social Psychology, 66 (2), 276-286.

Higgins, E. T., Tycocinski, O., & Vookles, J. (1990). Patterns of self-beliefs: The psychological significance of relations among the actual, ideal, ought, can, and future selves. Self-inference processes: The Ontario Symposium, Vol. 6. Ontario Symposium on Personality and Social Psychology. (J. M. Olson, & M. P. Zanna, Eds.), pp. 153-190.

Hilton, D. J. (1995). The social context of reasoning: Conversational inference and

rational judgment. Psychological Bulletin, 118(2), 248-271.

Hirst, D. E., Joyce, E. J., & Schadewald, M. S. (1994). Mental accounting and outcome contiguity in consumer-borrowing decisions. Organizational Behavior and Human Decision Processes, 58, 136-152.

Holyoak, K. J. & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. Journal of Experimental Psychology: General, 128, 3-31.

James, W. (1948). Psychology. New York: World. (Original work published 1890).

Kagel, J. H., & Levin, D. (1986). The winner's curse and public information in common value auctions. American Economic Review, 76, 894-920.

Kagel, J.H., & Roth, A.E. (Eds.) (1995). The Handbook of Experimental Economics. Princeton NJ:Princeton University Press.

Kahneman, D. (1991). Judgment and decision making:A personal view. Psychological Science, 2, 142-145.

Kahneman, D., Fredrickson, D.L., Schreiber, C.A., & Redelemeier, D.A. When more pain is preferred to less: Adding a better end. Psychological Science, 4, 401-405.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. Psychological Review, 93, 136-153.

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

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. Cognitive Psychology, 3, 430-454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. Psychological Review, 80, 237-251.



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

Kahneman, D., & Tversky, A. (1982). Psychology of preferences. Scientific American, 161-173.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. Psychological Review, 105, 582-596.

Kramer, R. M. (1995). Self-enhancing cognitions and organizational conflict. Working paper.

Kramer, R. M., Newton, E., and Pommerenke, P. (1991). Self-enhancement biases and negotiator judgment: Effects of self-esteem and mood. Organizational Behavior and Human Decision Processes.

Kuhn, T. S. (1970). The Structure of Scientific Revolutions (2<sup>nd</sup> edition). Chicago: University of Chicago Press.

Larrick, R. P., & Blount, S. (1997). The claiming effect: Why players are more generous in social dilemmas than in ultimatum games. Journal of Personality and Social Psychology, 723, 810-825.

Lazar, A. (1999). Deceiving oneself or self-deceived? On the formation of beliefs “uner Mind, in press.

Lecky, P. (1961). Self-consistency: A theory of personality. New York: McGraw-Hill.

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

Leven, S. J., & Levine, D.S. (1996). Multiattribute decision making in context: A dynamic neural network methodology. Cognitive Science, 20, 271-299.

- Lewin, K. (1935). A Dynamic Theory of Personality. New York:McGraw-Hill.
- Lewin, S. B. (1996). Economics and Psychology: Lessons for our own day from the early twentieth century. Journal of Economic Literature, 34, 1293-1323.
- Loewenstein, G. (1996). Out of control: Visceral influences on behavior. Organizational Behavior and Human Decision Processes, 65 (3), 272-292.
- Machina, M. J. (1982). 'Expected utility' analysis without the independence axiom. Econometrica , 50, 277-323.
- March, J. G., & Simon, H. A. (1958). Organizations. New York: John Wiley & Sons.
- Markman, A. B. & Medin, D. L. (1995). Similarity and alignment in choice. Organizational Behavior & Human Decision Processes, 63, 117-130.
- Markus, H., & Nuris, P. (1987). Possible selves. In K. M. Yardley, & T. M. Honess (Eds.), Self and identity: Psychosocial perspectives (pp. 157-172). New York: Wiley.
- Marr, D. (1970). Vision. San Francisco: W.H. Freeman.
- Mead, G. H. (1934). Mind, self, and society. Chicago: University of Chicago Press.
- Medin, D. L., Schwartz, H. C., Blok, S., & Birnbaum, L. (in press). The semantic side of decision making. Psychonomic Bulletin and Review.
- Medvec, V. H., Madey, S. F., and Gilovich, T. (1995). When less is more: Counterfactual thinking and satisfaction among Olympic medalists. Journal of Personality and Social Psychology, 69, 603-610.
- Messick, D. M. (1999). Alternative logics for decision making in social settings. Journal of Economic Behavior and Organizations, in press.
- Miller, N. E. (1944). Experimental studies of conflict. In J. Hunt (Ed.), Personality and behavioral disorders. Vol. 1 (pp. 431-405). New York: The Ronald Press.

Montgomery, H. (1983). Decision rules and the search for a dominance structure: Towards a process model of decision making. In P.C. Humphries, O. Svenson, and A. Vari, (Eds.), Analyzing and Aiding Decision Processes. Amsterdam: North-Holland.

Morris, M. W., & Larrick, R. P. (1995). When one cause casts doubt on another. A normative analysis of discounting in causal attribution. Psychological Review, 102-118.

Morris, M. W., Sim, D., & Girotto, V. (1998). Distinguishing sources of cooperation in the one round prisoner's dilemma: Evidence of cooperative decisions based on illusion of control. Working paper.

Murphy, G. L. & Medin, D.L. (1985). The role of theories in conceptual coherence. Psychological Review, 92, 289-316.

Neustadt, R. E., & May, E. R. (1986). Thinking in Time. NY: Free Press.

Nisbett, R. E., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment. Englewood Cliffs, NJ: Prentice-Hall.

Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment on property pricing decisions. Organizational Behavior & Human Decision Processes, 39(1), 84-97.

Ostrum, E. (1998). A behavioral approach to the rational choice theory of collective action. American Political Science Review, 92, 1-22.

Palmer, S. E., & Kimchi, R. (1986). The information processing approach to cognition. In T. J. Knapp & L. C. Robertson (Eds.), Approaches to cognition: Contrasts and controversies (pp. 37-77). Hillsdale, NJ: Erlbaum.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. Annual Review of Psychology, 43, 87-131.

- Piers, G., & Singer, M. B. (1971). Shame and guilt. New York: Norton.
- Rogers, C. R. (1961). On becoming a person. Boston: Houghton Mifflin.
- Roth, A. E. (1991). An economic approach to the study of bargaining. In M. H. Bazerman, Lewicki, R. J., & B. H. Sheppard (Eds.), Handbook of Negotiation Research: Research in Negotiation in Organizations, Vol. III, Greenwich, CT: JAI Press.
- Schafer, R. (1967). Ideals, the ego ideal, and the ideal self. In R. R. Hold (Eds.), Motives and thought: Psychoanalyst essays in honor of David Rapaport [Special issue]. Psychological Issues, 52(3), 131-174.
- Schelling, T. C. (1984). Choice and consequence: Perspectives of an errant economist. Cambridge, MA: Harvard University Press.
- Seligman, M. E. P. (1991). Learned Optimism. New York: A. A. Knopf.
- Sen, A. K. (1977). Rational fools: A critique of the behavioral foundations of economic theory. Philosophy and Public Affairs, 6(4), 317-345.
- Shafir, E. (1995). Compatibility in cognition. In J. R. Busemeyer, R. Hastie, & D.L. Medin, (Eds.), The Psychology of Learning and Motivation, Vol. 32. San Diego: Academic Press.
- Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. Cognitive Psychology, 24, 449-474.
- Shefrin, H., & Thaler, R. H. (1988). The behavioral life-circle hypothesis. Economic Inquiry, 26, 609-643.
- Sidgwich, H. (1874). The methods of ethics. London: Macmillan.
- Simon, H. A. (1957). Models of man. New York: Wiley.
- Slovic, P., & MacPhillamy, D. (1974). Dimensional commensurability and cue utilization in comparative judgment. Organizational Behavior and Human Performance, 11, 172-94.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. Review of Economic Studies, 23, 165-180.

Taylor, S. E. (1989). Positive illusions. NY: Basic Books.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective. Psychological Bulletin, 103, 193-210.

Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. Unpublished manuscript.

Thagard, P. (1989). Explanatory coherence. Behavioral and Brain Sciences, 12, 435-502.

Thaler, R. H. (1985). Using mental accounting in a theory of purchasing behavior. Marketing Science, 12-13.

Thaler, R. H., & Johnson, E. J. (1990) Gambling with the house money and trying to break even: The effects of prior outcomes and trying to break even: The effects of prior outcomes on risky choice. Management Science, 36, 643-660.

Thaler, R. H., & Shefrin, H. M. (1991). An economic theory of self control. Journal of Political Economy, 89, 392-406.

Thompson, L. (1998). The Mind and Heart of the Negotiator. New York: Prentice-Hall.

Tversky, A., & Kahneman, D. (1971). The belief in the "law of numbers." Psychological Bulletin, 76, 105-110.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 5, 207-232.

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

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions.

Journal of Business, 59, 251-284.

Tversky, A., & Sattath, S. (1979). Preference trees. Psychological Review, 95, 371-384.

Tyler, T., & Hastie, R. (1991). The social consequences of cognitive illusions. In M. H. Bazerman, R. J. Lewicki, & B. Sheppard (Eds.), Handbook of negotiation research: Research on negotiation in organizations. Greenwich, CT: JAI Press.

Usher, M. & Zakay, D. (1993). A neural network model for attribute-based decision processes. Cognitive Science, 17, 349-396.

Valley, K. L., Thompson, L. L., Gibbons, R., & Bazerman, M.H. (under review). Is Talk Really Cheap? Outperforming Equilibrium Models of Communication in Bargaining. Manuscript under review.

Wallsten, T. S. (Ed.). (1980). Cognitive processes in choice and decision behavior (pp. 263-274). Hillsdale, NJ: Erlbaum.

Walsh, V. (1994). Rationality as self-interest versus rationality as present aims. AEA Papers and Proceedings, 84 (2), 401-405.

Walsh, V. (1996). Rationality, allocation, and reproduction. Oxford: Clarendon Press.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. Psychological Bulletin, 116 (1), 117-142.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. Journal of Personality and Social Psychology, 60 (2), 181-192.

Winston, G. C. (1980). Addiction and backsliding: A theory of compulsive consumption. Journal of Economic Behavior and Organization, 1, 295-324.

Wylie, R. C. (1979). The self-concept. Lincoln: University of Nebraska Press.



Author Note

Douglas L. Medin, Department of Psychology; Max Bazerman, J.L. Kellogg Graduate School of Management.

This work was partially supported by NIMH grant MH55079 to the first author. Lance Rips, Scott Atran, Art Markman, Larry Birnbaum, and Ariela Lazar provided invaluable comments and suggestions.

Correspondence concerning this paper may be addressed to:

Douglas Medin  
Department of Psychology  
Northwestern University  
2029 Sheridan Road  
Evanston, IL 60208-2710  
[medin@nwu.edu](mailto:medin@nwu.edu)

or

Max Bazerman  
J.L. Kellogg Graduate School of Management  
2001 Sheridan Road  
Evanston, IL 60208-2001  
[Mbazer@nwu.edu](mailto:Mbazer@nwu.edu)



Table 1

Choice Option Used by Tversky and Kahneman (1986)

Options	Marbles (%)				
	White	Red	Green	Blue	Yellow
A	90	6	1	1	2
Outcome	\$0	Win \$45	Win \$45	Lose \$10	Lose \$15
B	90	6	1	1	2
Outcome	\$0	Win \$45	Win \$30	Lose \$15	Lose \$15
C	90	7	1		2
Outcome	\$0	Win \$45	Lose \$10		Lose \$15
D	90	6	1		3
Outcome	\$0	Win \$45	Win \$30		Lose \$15

## Footnotes

---

<sup>1</sup> Recursive decomposition is not to be equated with reductionism---different levels of description may be qualitatively different and higher-level descriptions may manifest emergent properties that derive from how the system is organized.

<sup>2</sup> While we find Marr's levels very useful, we see it as unfortunate that he used the term computational for his first stage. In many ways, the ordinary usage of the word *computational* is closer to Marr's algorithmic level.