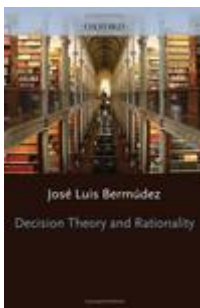


Rational Uses of Decision Theory

A review of



Decision Theory and Rationality

by José Luis Bermúdez

New York, NY: Oxford University Press, 2009. 189 pp. ISBN

978-0-19-954802-6. \$50.00



Reviewed by

[Gordon Pitz](#)

Most studies of decision making have at their core, implicitly or explicitly, the principles of Bayesian decision theory. Beliefs can be represented as probabilities and preferences as utilities. The theory can be used descriptively—to predict and explain people's choice behavior—or prescriptively—to assist people in making decisions or to evaluate their eventual choices.

The term *rationality* has also been widely used in discussions of decision making. It too has a prescriptive usage—defining what is rational—and a descriptive usage—determining to what degree people are rational. A common definition of *rational* considers it equivalent to making a choice that maximizes expected utility, as determined by decision theory. In *Decision Theory and Rationality*, José Luis Bermúdez provides an analysis of this interpretation. His conclusion is surprising and, on the face of it, potentially disruptive of an entire body of research on decision making.

Challenges and Contradictions

Bermúdez begins by pointing to three distinct functions that can be served by decision theory. The theory may be used to explain behavior (“Why did she do that?”). It may be used to evaluate someone’s analysis of a problem and that person’s eventual choice (“Was that a good thing to do?”). And it may be used to guide a person who seeks help in making a decision (“How can I help him achieve his goals?”).

Bermúdez also identifies three challenges that decision theory needs to overcome when used as a definition of rationality. Each challenge consists of inherent difficulties in determining what *rational* means within the framework of decision theory. He reviews strategies for coping with each challenge and evaluates the effectiveness of the strategies in addressing the three associated dimensions of rationality. He argues that while each strategy is appropriate for one or two of the dimensions, none can handle all three satisfactorily.

So far, so good. Decision theory serves three different purposes; in defining what is rational for each purpose, we confront three challenges. For each challenge there are three strategies for dealing with the challenge (presumably, the fact that there are three of everything is purely a coincidence). However, no set of strategies works satisfactorily for more than one purpose. The conclusion, then, is that the definition of rationality used for each application of decision theory must be different from the definition used for other purposes.

It takes Bermúdez a little over 160 pages to develop the argument to this point. In 10 pages more, he presents an apparent coup de grâce for decision theory as the criterion for rationality. He argues that a version of decision theory used for one purpose must necessarily be a theory for all three. A theory used for, say, description and explanation must be a theory for guiding decisions and for normative evaluation as well. We are left with an alarming proof by contradiction: No version of decision theory can serve all three purposes; any version that serves one purpose adequately must serve all three; therefore, no version of decision theory can serve any purpose adequately.

Challenges to the Definition of Rationality

The challenges presented by Bermúdez have been a source of disagreement for many years, and the literature on each is extensive. Anyone interested in either the descriptive or prescriptive use of decision theory will find his review to be useful, regardless of the conclusions one might derive.

Fundamental to decision theory is the assumption that a utility function can be used to describe a person’s preferences and values. There are at least three ways to interpret the

concept of utility, and therein lies the first challenge: Which interpretation should we adopt? Perhaps the simplest approach is an operational one: Utility is a representation of preferences as realized through a person's choices. An alternative is to consider utility to be a measure of preference involving more deliberation than is manifested by a simple choice. Finally, we might try to develop some external measure of the value that is assessed by utility.

Bermúdez considers the suitability of these three interpretations for the three dimensions of rationality. He concludes that only the first is suitable for the guidance dimension, only the second works for explanation and prediction, and only the third works for evaluation. That is, each dimension requires a different understanding of the core concept of utility.

The second challenge is one of great interest to psychologists. The large literature devoted to framing effects (see Kahneman & Tversky, 2000) shows how apparently irrelevant changes in the wording of a problem can lead to dramatically different decisions. A classic example is the Asian disease problem, in which treatment preferences depend on whether the problem specifies "lives saved" or "lives lost," even though one quantity can be inferred from the other. Under what conditions can we conclude that this change in preferences violates the principles of decision theory and is therefore irrational?

The principle being violated is the principle of invariance, which states that two propositions must have equal utility if they are materially equivalent. But in that simple principle lies a multitude of complexities that have bedeviled decision theorists. Simply stated, how do we determine that two propositions are materially equivalent? Does it really make no difference if we have been given information about lives saved or about lives lost?

Again, Bermúdez discusses three ways to address the issue. He finds that any of the strategies is adequate for guidance purposes. He argues that when decision theory is used for evaluation, only a strict adherence to the invariance principle, using an extensional interpretation of "equivalent events," is suitable for defining a rational choice. For explanation and prediction, the extensional interpretation will clearly not work (this is precisely the point made by studies of framing effects), and we must define equivalency in some other way.

The third challenge stems from inconsistencies in behavior over time. The inconsistencies emerge as a conflict between preferred plans for future choices and choices made when the time for action arrives. For example, I plan to avoid dessert with dinner tonight, but when dinner time comes I eat it anyway. Again Bermúdez presents three ways to determine whether an inconsistency is rational: two that work for the guidance dimension but not for evaluation or prediction, and one that works for evaluation and prediction but not for guidance.

What Does It All Mean?

When a philosopher presents a logical critique of a theoretical enterprise, scientists and practitioners tend to be skeptical. A psychologist's first response to Bermúdez would probably be empirical and pragmatic. If I can develop a testable theory that makes successful predictions, does it matter if the theory implies a definition of rationality that contains contradictions? If my guidance for decision makers leads to choices that are judged more satisfactory than they would be otherwise, does it matter that calling the behavior rational involves implicit contradictions?

Perhaps it does matter. In that case we need to separate out issues related to rationality as a concept and issues concerning decision theory itself. Sometimes it is not clear whether the problem Bermúdez sees is solely with the concept of rationality or with the use of decision theory to explain, guide, and evaluate behavior. If it is the former and all he is arguing is that we abandon use of the word *rational*, many decision theorists and psychologists will not object.

The slipperiness of the term *rationality* has been recognized for a long time (e.g., Brown, 1988). For many years psychologists wasted a great deal of effort and ink in debating whether humans are irrational. While one can still find vestiges of that argument in the literature (see Nickerson, 2008), attention has largely shifted to process-oriented explanations for behaviors that are less value laden.

The formal rules of decision theory still inspire respect, though. In spite of the body of research demonstrating departures from the prescriptions of decision theory, a number of authors still find the theory to be useful, supplemented if necessary with other assumptions (e.g., Gintis, 2009). In fact, Bermúdez provides a valuable assessment of the various strategies that might be adopted for using decision theory descriptively.

But if I adopt, say, Gintis's (2009) rational actor model to explain a person's behavior, using Bermúdez's second strategy for interpreting utilities, does this mean that I must adopt the same strategy when I use decision theory to help someone else deliberate over a difficult decision? I'm not convinced that it does.

And surely my use of decision theory for description or for guidance does not carry over to my evaluation of other people's choices. If I disagree with someone's decision, I can use decision theory as a language for discussing our disagreement, but I do not use it as an arbiter. Again, it's hard to see why the strategy I use for explaining someone's behavior must also be employed in settling our disagreements.

Part of my skeptical reaction to Bermúdez's argument is based on its inherently binary form. For a particular challenge to a particular dimension of rationality, a particular strategy is said to be successful or not successful. However, in many cases in which Bermúdez finds a strategy to be unsuccessful, I would prefer to call it "possibly less than ideal." He requires that a form of decision theory used for one purpose be a completely successful theory for

other purposes. But while I can see that a strategy used for, say, guidance might have implications for prediction and evaluation, I am not convinced that it must be fully satisfactory in every respect. If one takes a more nuanced view of these issues, the proof by contradiction collapses.

Bermúdez makes an important contribution to an area that has already generated a great deal of debate. His description of the challenges to decision theory and the strategies for addressing them is thought provoking. If he succeeds in making psychologists and others more reluctant to use the word *rationality*, it will be no great loss. I think it unlikely, though, that his arguments will restrict the use of decision theory in some form as a framework for understanding, guiding, and evaluating choice behavior.

References

- Brown, H. I. (1988). *Rationality*. New York, NY: Routledge. [PsycINFO](#)
- Gintis, H. (2009). *The bounds of reason: Game theory and the unification of the behavioral sciences*. Princeton, NJ: Princeton University Press. [PsycINFO](#)
- Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. New York, NY: Cambridge University Press. [PsycINFO](#)
- Nickerson, R. S. (2008). *Aspects of rationality: Reflections on what it means to be rational and whether we are*. New York, NY: Psychology Press. [PsycINFO](#)
-