



“Summary of article by Daniel Kahneman: New Challenges to the Rationality Assumption” in Frontier Issues in Economic Thought, Volume 3: Human Well-Being and Economic Goals. Island Press: Washington DC, 1997. pp. 177-180

Social Science Library: Frontier Thinking in Sustainable Development and Human Well-being

“Summary of article by Daniel Kahneman: New Challenges to the Rationality Assumption”

"In the domain of social policy, the rationality assumption supports the position that it is not necessary to protect people against the consequences of their choices. The status of this assumption is therefore a matter of considerable interest." (18)

In economic theory it is assumed that rational persons are utility maximizers. However, new empirical evidence challenges the assumption that choices, in fact, maximize utility. Two findings stand out: first, individuals cannot be assumed to be infallible forecasters of future tastes and secondly, they are often unable to accurately evaluate past experience. This paper argues that the traditional hallmark of rationality, i.e., consistency among preferences, is insufficient as a criterion of rationality. Rationality can and should be assessed using substantive criteria, such as experienced utility, that are independent of the system of preferences.

For several decades, debates concerning rationality have focused on the paradoxical implications of the consistency standard. A number of researchers have shown that certain pairs of “reasonable” preferences lead to paradox when combined with the axioms of expected utility theory. Resolution of these paradoxes became the focus of decision theory, reinforcing consistency as the criterion of rationality.

MULTIPLE NOTIONS OF UTILITY

In classical usage, "utility" meant experienced utility, the hedonic experience of an outcome. Early utilitarian theorists such as Jeremy Bentham focused on the pleasure and pain that accompanied outcomes. Subjective mental states, however, lost their status as a legitimate area of study as behaviorism gained currency in the social sciences during the twentieth century. Today, outcomes are evaluated in terms of their desirability rather than the affect derived from them. Two assumptions prevail in the current discourse on utility: people reveal their preferences in the choices they make and people enjoy later what they want now. As a result, “utility” in modern usage means decision utility, the weight assigned to an outcome of a decision.

Recent work in psychology indicates an empirical basis for revitalizing Bentham's notion of utility. Support is proffered for an additional substantive criterion of rationality -- how well do people maximize experienced utility. Substantive criteria of rationality can be assessed in three

areas: decision utility, predicted utility, and the relation between real-time and retrospective utility.

DECISION UTILITY

Standard economic assessments of rationality omit key features of the human decision-making process. Research findings indicate three types of deficiency: utility is assigned to gains and losses, not to absolute levels (the main carriers of utility are events, not states); losses seem larger than corresponding gains (loss aversion); and, the same objective outcomes can be evaluated as gains or losses, depending on the frame of reference (framing effect). A set of ingenious experiments reveals two psychological effects that demonstrate the complexity involved in weighting outcomes. First, the framing effect illustrates that wealth is valued in relative rather than absolute terms and that inconsequential differences in the formulation of a problem lead to diverse preferences. Consider the two problems below:

Problem 1: Assume yourself richer by \$300 than you are today. You have to choose either a sure gain of \$100 or a gamble: 50 percent chance to gain \$200 and 50 percent chance to gain nothing.
Problem 2: Assume yourself richer by \$500 than you are today. You have to choose either a sure loss of \$100 or a gamble: 50 percent chance to lose nothing and 50 percent chance to lose \$200.

Both problems offer the subject a choice between a sure gain of \$400 or a gamble with equal chances to increase current wealth by \$300 or \$500. Mainstream accounts imply that rational people should treat these problems the same, since the amounts of wealth are the same. In fact, people treat these problems differently, favoring the sure thing in problem one and the gamble in problem two.

Second, the endowment effect demonstrates loss aversion and suggests that people are extremely myopic when it comes to weighting future outcomes. In one experiment, some subjects were given a coffee mug and asked the amount of money for which they would be willing to exchange it. Other subjects were offered a choice between owning a mug and receiving an amount of money that they thought the mug was worth. In an important sense, both groups of subjects were offered the same choice: leave the experiment with a new mug or extra pocket money. Owners in the first group evaluated the difference between having a mug and not having one as a loss, while choosers in the second group evaluated the difference as a gain. Consistent with the assumption of loss aversion, mug owners ascribed a much higher average cash value (\$7.12 in one experiment) than for choosers (\$3.50). Although all subjects chose between the same long-term states -- "own this mug" or "not-own this mug," subjects based their decisions on evaluations of short-term transitions -- "receive a mug" or "give up your mug." Such myopia appears to be linked with the emotions involved in making transitions. The endowment effect has implications for both logical and substantive criteria of rationality: given different representations of the same problem people reveal inconsistent preferences and individuals are led to inferior outcomes when they myopically evaluate future consequences.

PREDICTED UTILITY: DO PEOPLE KNOW WHAT THEY WILL LIKE?

Experiments show that people are often unable to accurately predict hedonic responses to future stimuli. In one study of the endowment effect, subjects were shown a coffee mug. They were then asked to imagine being given a replica and given the opportunity to either keep it or trade it for money. After subjects decided on an exchange rate for the imaginary mug, they were actually given a mug, and then asked how much money they would be willing to trade it for. The mean selling price after receiving the mug (\$4.89) was much higher than that predicted (\$3.73), which suggests that people are unable to anticipate that possession of a mug would induce a reluctance to give it up. A different experiment indicates that people are unable to accurately predict their tastes. In this study, most subjects, after tasting a spoonful of plain low-fat yogurt, predicted that their hedonic response to a six ounce helping on the next day would be the same. Not surprisingly, the larger helping was a much worse experience than anticipated. "The data provided no indication that individuals were able to predict the development of their tastes more accurately than they could predict the hedonic changes of a randomly selected stranger." (27)

RETROSPECTIVE UTILITY: DO PEOPLE KNOW WHAT THEY HAVE LIKED?

In order to know what will be liked in the future, it is invaluable to know what has been liked in the past. However, individuals employ a surprisingly fallible process to represent past experiences. Global evaluations of past experiences require two operations: recollecting the momentary experiences that constitute a given episode and combining the affect associated with these moments. Counter to normative intuitions, experiments strongly suggest that these operations, recollecting experience and combining affect, are performed without adding utilities. In fact, studies support two alternative empirical generalizations: a peak and end rule, which proposes that global evaluations of past episodes are formed by weighting the experience of most extreme affect and the affect experienced during the final moments of the episode; and, duration neglect, which considers global evaluation of total affect unaffected by an episode's duration. For example, in one study colonoscopy patients reported pain every sixty seconds in procedures that took anywhere from four to sixty-nine minutes, yet both patients and attending physicians made retrospective evaluations that reflected the intensity of pain at its worst and the intensity of discomfort during the last few minutes of the procedure.

CONCLUSION

Adding substantive criteria to the current logical standard of consistency makes sense on a number of levels. Substantive analyses permit a more demanding definition of rationality than the consistency standard commonly invoked in economic discourse. Foolish decisions and myopic choices need not be considered rational just because they are part of a system of coherent preferences. A better understanding of hedonics could reveal a great deal about the welfare consequences of institutions and identify which skills derive the most experienced utility from outcomes. "The time has perhaps come to set aside the overly general question of whether or not people are rational...What are the conditions under which the assumption of rationality can be retained as a useful approximation? Where the assumption of rationality must be given up, what are the most important ways in which people fail to maximize their outcomes?" (35)