Paweł Rybacki  
Perspectives: Assignment 2  
November 10, 2021

Cognitive scientists, such as Jonathan Haidt, the author of “The Happiness Hypothesis: Finding Modern Truth in Ancient Wisdom,” challenge some fundamental assumptions regarding human rationality prevalent among many social scientists. The insights from psychology and neuroscience help economists, sociologists, political scientists, and others not only rethink the assumptions of the rational choice theory, but also put the theory into a proper perspective and provide complementary explanations. This essay is an attempt to critique Haidt’s (2006) from the perspective of the rational choice theory by showing that while his argument somewhat underestimates the breadth of the theory’s applicability, it also helps us understand some of its limitations and suggests how to address them.

Haidt (2006) argues that human minds are more complex than contemporary social scientists tend to think about them. The observation that humans exhibit rational behavior inconsistently (or “keep doing … stupid things”) motivates his critique of the rational choice approach, which does not “adequately explain weakness of the will” (Haidt, 2006, p. 3). To illustrate his angle of critique, the author contrasts the modern way of thinking about mind as a mechanical process with ancient metaphors that compare the rational part to a rider and the irrational part to an animal, and he considers the latter more accurate (Haidt, 2006, p. 3).

To support the argument that humans do not think entirely rationally, controllably, or predictably, he presents evidence for four ways in which our minds can be divided into conflicting parts (mind/body, left/right, new/old, controlled/automatic) and goes on to present more general arguments from failures of self-control, moral intrusions, the irrationality of human moral or aesthetic judgement (Haidt, 2006, Chapter 1). Haidt concludes that the mind is too often equated with the rational part of the whole and calls for more attention to the unconscious part of the human mind (Haidt, 2006, p. 22).

Haidt’s (2006) proposition that human minds are “loose confederations of parts” (2006, p. 22) directly questions voices such as Becker’s claim that “human behavior is not compartmentalized, sometimes based on maximizing, sometimes not” (1976, p. 14). For Haidt (2006), maximizing behavior would only be applicable to the rational part of the human mind, but not to the intuitive or emotional parts. In this light, some classical articles of economics no longer appear as complete analyses of individual choices. Becker (1968), for example, models decisions to engage in criminal activity as an optimization problem given offenders’ potential individual costs, benefits, and risk preferences, but he makes no mention of the possibility that feelings, emotions, or affect can distort the decision-making process.

Another interesting contradiction to the rational choice theory is Haidt’s claim that “[r]eason and emotion must both work together to create intelligent behavior, but emotion … does most of the work.” (Haidt, 2006, p. 13). He supports this with a lab experiment case, in which people deprived of emotions as a result of brain damage—as close to *homo economicus* as real persons can get—faced difficulties in making decisions and “their lives f[e]ll apart” (Haidt, 2006, p. 12). However, examples of this kind do not necessarily prove the rational choice theory wrong under the qualification that *average* human behavior, measured with the models, can be perfectly rational, even if an *individual* human behavior cannot. Moreover, perhaps emotions and automatic processes drive most of human behaviors in daily lives; yet, at the same time, calculated rational behaviors could still drive most market interactions. Nonetheless, the primacy of emotions over rational calculation makes sense in political contexts such as revolutions. Purely rational agents would not be willing to “selflessly to gamble with their lives,” which is rather a behavior resulting from strong frustration. Alternatively, the rational utility function can be properly modified, like in Kuran (1991, p. 14), to incorporate the emotional component.

One of the least effective arguments against human rationality in Haidt (2006) is when the author likens the justification of intuition with confabulation.[[1]](#footnote-1) He refers to the observation that, when people try to explain their moral or aesthetic positions, they often have difficulty indicating the core underlying reason. Thus, while offering partial explanations, they rarely change their minds on morality or aesthetics even if the explanations are directly disproved. Haidt, therefore, attributes the moral and aesthetic judgements to automatic, or animal-like processes, that our conscious and controlled processes attempt to defend (Haidt, 2006, p. 21-22). Whereas this could be, at first glance, a compelling case against the consistency of human rationality, the author does not present further evidence to refute possible rational capabilities of human subconsciousness. Perhaps there are some rational processes happening in the background, which humans cannot readily recall or articulate, but which involved the expressed partial explanations? Regardless of the answer, many automatic processes seem predictable at the individual-level, which gives rise to stable preferences embraced by the rational choice framework.

An important limitation to the scope of Haidt’s (2006) critique of rational choice theory is the fact that the weakness of will may not pertain to groups equally well as to individuals, while rational choice applies to both individual and collective entities. Becker (1968), for instance, uses the rational choice framework not only to individuals considering illegal behavior, but also the state as a rational agent calculating its policy to achieve a social optimum. Similarly, Ostrom (1990) uses rational choice reasoning to explain how different organizations govern the commons. Haidt’s (2006) argument is not easily applicable to these cases. It would seem absurd to assume that a collective brain would experience a division between its left and right parts, for example, as human minds do not simply “add up”. Moreover, firms make decisions just within the scope of business activity rather than all aspects of life, which allows the decisions to be better calculated using actual cost-benefit analyses. And while some companies may suffer from leaders with self-control issues, competitive reality quickly clears markets from irrational actors.

In the light of the preceding discussion, it is crucial to realize that the definitions of rationality differ by much and overlap by little across the approaches of Haidt and of the rational choice theory.[[2]](#footnote-2) Because of this, they seem to explain the human behavior complementarily. As opposed to associating rationality with “conscious verbal thinking,” like in Haidt (2006, 22), the rational choice seems to assume the occurrence of an economic calculus regardless of how conscious or verbal it is—as Becker clarifies, “the economic approach does not assume that decisions units are necessarily conscious of their efforts to maximize or can verbalize … reasons for the systematic patterns in their behavior” (1976, p. 7). Rational choice theorists, therefore, seem exclusively interested in human behavior as it is observed externally; not in internal mechanisms behind mathematically predictable behavior. From this angle, Haidt’s description of the mechanics of internal tensions of mind appears to present an anatomy behind what social scientists externally observe as rational choice optimization.

For instance, if increased police presence deters potential offenders through the *feeling* of anxiety induced by their liminal systems, even such a mechanism behind lower willingness to commit crime still counts as a *rational* *response* to the increased perceived probability of bearing the costs of the action under Becker’s (1968) framework. Another example are predictable failures of self-control, such as when the preference to eat a lot of chocolate today rather than remain slim by the end of month, which may be incorporated in individual utility, regardless of whether the individual believes that this unhealthy preference is “rational” or not.

Ultimately, then, utility could be understood as a generalized notion of happiness, which is a feeling; as such, the goal of rational calculation could be to maximize the positive feelings. Thus, while Haidt’s insights do not seem to be an ultimate challenge to the economic approach as outlined in Becker (1976), they may enhance the way economists understand their assumptions, help redefine “rationality” and “utility,” and highlight the importance of behavioral models.

## Bibliography

Gary S. Becker. (1968). “Crime and Punishment: An Economic Approach.” Journal of Political Economy , Mar. - Apr., 1968, Vol. 76, No. 2 (Mar. - Apr., 1968), pp. 169-217. The University of Chicago Press.

Gary S. Becker. (1976). *The Economic Approach to Human Behavior*. The University of Chicago Press.

Jonathan Haidt. (2006). *The happiness hypothesis: Finding modern truth in ancient wisdom.* Basic Books.

Timur Kuran. (1991). “Now Out of Never: The Element of Surprise in the East European Revolution of 1989.” World Politics , Oct., 1991, Vol. 44, No. 1 (Oct., 1991), pp. 7-48 Cambridge University Press.

Elinor Ostrom. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge University Press.

1. The phenomenon whereby “people will readily fabricate reasons to explain their own behavior” (Haidt, 2006, p. 8). [↑](#footnote-ref-1)
2. To restate this for further clarification: Haidt’s critique of rational choice is certainly accurate if “rational” only refers to reason, rider, or the controlled behavior. However, “rational” seems to have a different meaning in “rational choice.” The interchangeably used phrases of “rational choice” and “optimal choice” tend to refer to all the processes in the brain taken together, as long as all of their goals are incorporated into the utility maximization function. [↑](#footnote-ref-2)