Kahneman and Tversky and the making of behavioral economics
Heukelom, F.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
6. Building and defining behavioral economics

1. Who are we?

George Loewenstein, a prominent behavioral economist, recalls that

In 1994, when Thaler, Camerer, Rabin, Prelec and I spent the year at the Center for Advanced Study in the Behavioral Sciences, we had a meeting to make a kind of final decision about what to call what we were doing. Remarkably, at that time, the name behavioral economics was not yet well established. I actually advocated ‘psychological economics,’ and Thaler was strong on behavioral economics. I'm kind of glad that he prevailed; I think it's a better, catchier, label, although it creates confusion due to association with Behaviorism. [email Loewenstein to author, June 16, 2008]

It was no accident that Richard Thaler gave the new field its name. Thaler’s behavioral finance had developed in the 1980s as a sub-field in financial economics, but towards the late 1980s and early 1990s Thaler realized that behavioral finance could be broadened to include other violations of neoclassical economics. One example is the anomaly paper of Thaler and Loewenstein [Loewenstein and Thaler (1989)] discussed in the previous chapter. It started with the violation of rational intertemporal choice in financial markets, but subsequently also applied the theory to subjects’ behavior in the dermatologist’s office. As a consequence of this broadening of behavioral finance, more economists and non-economists began to join Thaler’s research program. The central argument was that neoclassical economics was descriptively wrong in every domain to which it was and could be applied. This development culminated in the creation of the field of behavioral economics in 1994, as Loewenstein recalls in the above quote.

During the 1990s and 2000s, Thaler and other behavioral economists expanded behavioral economics from a small research program focused on violations of the neoclassical theory in financial economics to a dominant research program that looked for inspiration beyond behavioral decision research to a range of scientific disciplines and methods. To date, they have investigated violations in every aspect of life and from the early 2000s onwards they have developed their own stance on policy advice under the rubric of libertarian, light, or asymmetric paternalism. As a result of
this expansion, behavioral economists came to realize how behavioral economics relates to neighboring fields. Thus, in the late 1990s and 2000s, behavioral economics started to distinguish itself more explicitly from neighboring fields, such as experimental economics and psychology. All of these developments helped to expand behavioral economics into a broad and stable economic research program in which the influence of Daniel Kahneman and Amos Tversky relatively declined. The two psychologists remained the iconic founding fathers, but the omnipresent influence they had achieved in Thaler’s behavioral finance in the 1980s no longer existed. Kahneman and Tversky’s normative – descriptive distinction remained the methodological basis of behavioral economics, but the labels were changed into full rationality versus less-than-fully, quasi, or, eventually, bounded rationality. This change of terminology made behavioral economists redefine the history and subject matter of their field, no longer claiming that only Kahneman and Tversky were their founding fathers, but Herbert Simon as well.

This chapter presents the story of how behavioral economics was built and redefined in the 1990s and 2000s. Of course, the literature in behavioral economics is so vast that it is impossible to do justice to the wide variety of research within the scope of one chapter. I shall, therefore, proceed by discussing the examples which can serve to illustrate the move to expand and redefine the field in crucial ways. Research into intertemporal choice and emerging preferences challenged standard assumptions in neoclassical theory about the time consistency of choice, and the given character of preferences. Using Kahneman and Tversky’s distinction between a normative benchmark and descriptive deviations, this research moved in some instances even so far as to challenge the normative benchmark of rational choice itself. As we shall see, this last move was for some behavioral economists a ‘bridge too far.’ The research into preference formation shows how a split emerged between those economists who did not accept the normative benchmark and who moved into the direction of research such as that of psychologist Gerd Gigerenzer (under the label of “ecological rationality”), and those economists who kept adhering to the normative-descriptive distinction, albeit in modified form. Both strands claimed, as later will be seen, Simon as their founding father, but for different reasons. We will see that behavioral economists used the terminology of bounded rationality to explore the rapidly growing field of neuroeconomics. Here, we see the independent influence that using new tools of research had on economists’ practices as well.
Behavioral economists gradually built their program into a stable and well-defined mainstream economic program. In this process, the main question was how to construct the descriptive theory of human decision behavior. To answer that question, behavioral economists explored a range of different scientific disciplines and methods. At the same time, however, they remained faithful and always came back to the normative – descriptive framework that had originally been introduced by Kahneman and Tversky. This conceptual core determined how behavioral economists understood the economic world, it determined the welfare implications they drew, and finally, it determined how they pulled back when their explorations diverged too far from this conceptual core. The new terminology of rationality lay the foundation for discussing behavioral economists’ new paternalistic stance on economic policy advice which developed from the early 2000s. The building of behavioral economics from the mid 1990s onwards led behavioral economists to distinguish themselves more clearly from psychology and experimental economics. Redefining their own field in the process, behavioral economics is now one of the most thriving and innovative branches of economics as a discipline.

2. Building behavioral economics

2.1 Intertemporal choice and the dual system approach

Behavioral economic research on intertemporal choice started in the early 1990s and culminated in the behavioral economic research based on the two-systems approach. The intertemporal choice and two-systems approach literature is illustrative for a number of reasons. First, it was chronologically the first major theme behavioral economists focused on and it continues to this day to be an important topic in behavioral economics. Second, it illustrates how behavioral economists incorporated Kahneman and Tversky’s work, and in particular their normative – descriptive distinction into economics. During this process behavioral economists sometimes would venture so far as to also question the normative benchmark, but in the end they always pulled back to the conceptual foundation as laid out by Kahneman and Tversky. Finally, it illustrates how behavioral economists developed a theoretical framework that could be applied to any economic problem and was compatible with important developments in neuroscience and the cognitive sciences in general.

The two most prominent behavioral economists who have worked on intertemporal choice have been George Loewenstein and David Laibson. Loewenstein
finished his PhD at Yale in 1985 and published his first article in 1987. From his first publication onwards he has been a strong proponent of more psychology in economics, but initially he was hardly influenced by the work of Kahneman and Tversky. Instead, an important theoretical influence came from the work of Jon Elster, with whom he wrote several articles and edited a book for the Russell Sage Foundation called *Choice over Time* (1992). Loewenstein has published a number of articles on the history of psychological and economic explanations of intertemporal choice and utility, revealing an extensive knowledge of the history of economic thought [e.g. Loewenstein (1992), Elster and Loewenstein (1992), Frederick, Loewenstein, and O’Donoghue (2002), Angner and Loewenstein (forthcoming a,b)].

Laibson finished his PhD at MIT in 1994 and started his academic career at Harvard University that same year. At MIT and Harvard he has focused on violations of the traditional economic idea of exponential discounting. His articles are a mixture of experimentally corroborating this phenomenon, building mathematical economic models that account for the observed systematic deviations, and investigating the psychological and neurobiological substrates of the observed behavior.

For Loewenstein the problem of the well-known exponential discounting utility (DU) model was not just that individuals discount hyperbolically, but it goes even further. For instance, individuals can be shown to sometimes use a negative discount rate [Loewenstein and Prelec (1991)]. If individuals prefer an increasing real-wage over a constant real-wage, even when the present value of the latter is higher than the former, they effectively employ a negative discount rate. Perhaps even more challenging for received economic theory was that individuals’ intertemporal choices could be shown to be fundamentally inconsistent [e.g. Prelec and Loewenstein (1997)]. People who prefer *A now* over *B now* also prefer *A in one month over B in two months*. However, at the same time they also prefer *B in one month and A in two months* over *A in one month and B in two months*. In other words, when faced with an intertemporal choice individuals like to save the best for last, which is in fundamental disagreement with economic theory. Another, by now famous descriptive falsification of the DU model, is the research on New York City cab drivers who judge their income “one day at a time” [Camerer, Babock, Loewenstein, and Thaler (1997)].

---

40 See for a discussion on Elster’s work on decision making e.g. Davis (2003).

41 Here, Loewenstein referred to Friedman and Savage (1948, 1952) who face an analogous problem in explaining both gambling and insurance behavior.
The DU model fails not only descriptively, but also normatively, Loewenstein argued. For instance, there does not seem to be a good reason to suppose that somebody who is indifferent towards oranges and apples today should also be indifferent towards 1) apples today, oranges tomorrow, and apples the day after and 2) apples three days in a row. Loewenstein argued that there is little normative and descriptive reason for holding on to the DU model, despite its aesthetic merits of mathematical simplicity and consistency. This conclusion produced tension in Loewenstein’s work. On the one hand, he was an early proponent of Thaler’s behavioral finance, and a founding member of behavioral economics. But on the other hand, he concluded that economic theory might also be problematic as a normative benchmark.

Not surprisingly, then, Loewenstein was ambivalent about how to proceed. A number of publications show fundamental problems with the DU model both descriptively and normatively [e.g. Loewenstein (1992, 1999), Loewenstein and Prelec (1991), Prelec and Loewenstein (1997)]. But on different occasions Loewenstein also tried to extend the DU model as a descriptive theory which could fulfill its role while maintaining the normative benchmark. For instance, he built a mathematical model that can accommodate observed behavior while maintaining the normative benchmark as a limiting case [see e.g. Loewenstein and Prelec (1992)]. The discount factor is generalized to $1/(1+at)$, where $a$ can be exogenously given or determined by another function. Loewenstein also turned his attention to neuroscience as a possible means to a solution [see e.g. McClure, Laibson, Loewenstein, and Cohen (2004)].

Loewenstein’s explorations led him to doubt the normative – descriptive distinction of behavioral economics in the 1990s, but after a while he pulled back from contesting the normative benchmark and made his work compatible with Kahneman and Tversky’s approach. In the 2000s, Loewenstein acknowledged that his work had been stimulated “by the existence of a strong normative benchmark, expected utility theory,” that behavioral theory in economics and psychology had advanced, and that “both psychologists and economists have made important theoretical and empirical contributions” [Loewenstein, Weber, Hsee, and Welch (2001), p.367]. Moreover, he added that in this area the “convergence in the theoretical perspectives of psychologists and economists […] has been greater than for
any other topic of mutual interest in the two disciplines [Loewenstein, Weber, Hsee, and Welch (2001), p.367]. This use of Kahneman and Tversky’s normative and descriptive in behavioral economics will be discussed in more detail below in Section 2.3.

A similar tension can be observed in the work of Laibson. In Laibson (1997), “Golden Eggs and Hyperbolic Discounting,” he built a mathematical model of agents with hyperbolic discount functions that could explain a myriad of dynamically inconsistent individual preferences observed in experiments. “Golden Eggs” referred to the traditional rational economic individual decision model. In Laibson’s model, the individual was faced with an “imperfect commitment technology,” such as a retirement plan, which required that it be initiated one period before it started to work. Together with the hyperbolic discount function this model “predicted” that individuals’ consumption would closely track the progress of their income, but that with the “imperfect commitment technology” individuals were capable of correcting their hyperbolic discount functions by committing themselves in advance to their desired savings behavior. Because the imperfect commitment technology required individuals to commit themselves in advance, the far-sighted, rational planner effectively constrained the temptation to be immediately gratified once the money actually arrived.

Ipsa facto the model predicted that with “financial innovation” savings rates would go down because commitment technology no longer needed to be started up a period in advance. According to Laibson, this provided an explanation for the ongoing decline in U.S. saving rates. “Financial innovation” should be interpreted broadly here. It not only comprises new saving plans at banks, but also changes in “social commitment devices” such as marriage, work and friendship. The idea, which went back to Laibson’s dissertation, was that a decrease in the strength of the structure and/or duration of long-term social commitments increased the probability of acting according to the short term hyperbolic discount function.

Furthermore, Laibson also showed that the dynamically inconsistent preferences resulting from hyperbolic discounting might lead to a welfare reduction following financial innovation. Financial innovations allowed individuals’ short term hyperbolic discount functions to override their long term rational discount functions. Under certain conditions, the result may, from a rational, long-term perspective, be a reduction in welfare.
Laibson’s ‘Golden Eggs’ article is a typical 1990s contribution to behavioral economics. It made productive use of Kahneman and Tversky’s distinction between the normative and the descriptive by reinterpreting this in terms of a “far-sighted” and a “myopic” planner. Whereas in the 1980s Thaler would have used accessible language and would have referred to psychology to explain and solve the problem, Laibson presented the problem in a formalistic economic language and did not make any reference to psychology. Behind this emphasis on economics, however, Laibson was fully committed to the normative – descriptive distinction of Kahneman and Tversky. Golden eggs and hyperbolic discounting are the normative and descriptive rephrased in economic language. However, Laibson also added something to this framework, namely the idea that in the economy individuals might have the possibility of controlling their deviating behavior by means of commitment technologies. This idea of commitment technologies extended the Kahneman and Tversky framework. It suggested that individuals might be both deviating from the normative benchmark, while at the same time helped to explain how they deviate from the norm. Further, it suggested that it is not so much the scientists that need to find ways to correct individuals’ deviating behavior, but the individuals themselves.

Harris and Laibson (2001), “Dynamic Choices of Hyperbolic Consumers,” elaborated further on the idea of hyperbolic discounting. It tried to link the short term hyperbolic discounting with the long term (rational) exponential discounting and showed how individuals act who try to prevent their own future overconsumption. The paper started with the traditional discounting function for individuals, and replaced the constant discount factor $\delta$ with an “effective discount factor.” This effective discount factor consisted of the sum of two components, the “long-run discount factor $\delta$” and the “short-run discount factor $\beta\delta$,” where hyperbolic discounting implied $\beta < 1$. The traditional discount factor was hence explicitly decomposed into a long-run, exponential component and a short-run, hyperbolic component. The assumption was that individuals, faced with “stochastic income” and a “borrowing constraint,” anticipate their future inclination to hyperbolically discount (and thus to overconsume), and that they want to act against it. Hyperbolic discounting was thus explained as resulting from a strategic game with future selves.
Since $\beta < 1$, the effective discount factor is negatively related to the future marginal propensity to consume (MPC). To gain intuition for this effect, consider a consumer at time 0 who is thinking about saving a marginal dollar for the future. We assume that this consumer acts strategically in an *intrapersonal* game where the players are temporally situated “selves.” The consumer at time zero – ‘self 0’ - expects future selves to overconsume relative to the consumption rate that self 0 prefers those future selves to implement. Hence, on the equilibrium path, self 0 values marginal saving more than marginal consumption at any future time period. From self 0’s perspective therefore, it matters how a marginal unit of wealth at time period 1 will be divided between savings and consumption by self 1. Self 1’s MPC determines this division. Since self 0 values marginal saving more than marginal consumption at time period 1, self 0 values the future less the higher the expected MPC at time period 1. [Harris and Laibson (2001), p.936, emphasis in the original]

In equilibrium, self 0 would reduce his or her savings rate (and thus his or her future income) to the point where his or her preference for self 1’s savings rate would be equal to self 1’s actual savings rate. In other words, because individuals knew they would discount hyperbolically in the future, they would also discount hyperbolically now. In equilibrium the two selves maximize the combination of their preferences. The effective discount rate was a function of a “time preference” (the difference between preferences of self 0 and self 1) and an anticipation of future MPC.

Harris and Laibson (2001) also demonstrates how Laibson retreated from the idea that individuals can influence their own behavior through commitment. Instead, the behavior displayed was merely the result of conflicting selves, none of whom could be controlled by the other. Like Loewenstein, Laibson struggled between exploring new directions and remaining committed to a distinction between the norm and its imperfect realization.

The research on intertemporal choice illustrates how behavioral economists extended Thaler’s behavioral finance. First a piece of standard neoclassical economics was examined; in this case the assumption of exponential discounting. Subsequently, it was shown that this piece of the neoclassical theory failed descriptively. In the next step, the piece of neoclassical economics as descriptive theory was adjusted to be
compatible again with the empirical facts, the same approach as used in Kahneman and Tversky’s prospect theory.

The behavioral economists’ way of dealing with intertemporal choice described human behavior as the outcome of two systems or processes striving for dominance. Different labels appear in the behavioral economic literature for these two systems: reasoning vs. intuition [e.g. Kahneman (2003)], rationality vs. emotion [e.g. Shefrin and Thaler (1988), van Winden (2007), Ben-Shakar, Bornstein, Hopfensitz, and van Winden (2007)] and cognitive vs. affective [e.g. Camerer, Loewenstein and Prelec (2005)] are the most prominent. Understanding human behavior as the outcome of conflict between different motives has a long and rich history, going back to the philosophy of Plato and Aristotle and to Homer’s Ulysses tying himself to the mast so he could hear the Sirens sing [Davis (2003), pp.63-80]. Behavioral economists’ re-creation of neoclassical economics’ understanding of individual behavior in terms of two souls inhabiting one body is therefore a recent development in a long history.

Some behavioral economists have linked this dual system solution to research in neuroscience and neurobiology, thus contributing to the creation of a new sub-field called neuroeconomics.\footnote{The label neuroeconomics is used by at least two groups of scientists [see e.g. Vromen (2007), Ross (2008)].} This literature maintained the normative – descriptive distinction, but nevertheless slightly re-interpreted the distinction by supposing that the two sides of the distinction represent two sides of human behavior. In other words, the normative was reduced from something external to the individual to one of two faculties innate to human nature that strive for dominance. An illustrative example is McClure, Laibson, Loewenstein, and Cohen (2004), “Separate Neural Systems Value Immediate and Delayed Monetary Rewards.” The research described in the article sought and found evidence for neurobiological substrates for the two components of the effective discount factor as described above. When faced with delayed monetary rewards while lying in an MRI-scanner, subjects’ brains showed peaks of activity in the parts of the brain associated with rational behavior (in this case the lateral prefrontal cortex and the posterior parietal cortex); when faced with immediate rewards, the limbic system associated with the midbrain dopamine center was especially active.
The authors took their findings as evidence that, when faced with the choice between an immediate reward and a higher future (thus ‘expected’) reward, the two parts of the brain strive for dominance. The limbic system was especially sensitive to immediate rewards and signaled choosing the immediate reward. The prefrontal cortex was more sensitive to the higher expected pay-off and signaled choosing the delayed reward. In other words, the experimental results were taken as evidence for the postulated difference between the short-term hyperbolic discounting, and the long-term rational exponential discounting. The normative long run rational system strives for dominance with the short term affective system. When the short run system is affecting the outcome, the resulting behavior will be observed and classified as systematically deviating from the norm. As such, the neuroeconomic research conducted by behavioral economists at once maintained the link with the normative – descriptive core of Kahneman and Tversky’s work, while at the same time constructing a link with neuroscience.

Based on this and other research, in the 2000s behavioral economists have increasingly argued that the neuroscientific framework should be adopted as a basis for investigating individual (economic) behavior. The recurring argument in the neuroeconomic research has been that in economic decision making the individual’s rational system tries to make the rational decision, but, alas, is often overridden by a strong and dominant affective system. Intertemporal choice has provided a good example. When people need to plan how to divide their income between consumption and spending at some point in the future, their affective system is not very much involved and the rational system will decide on a rational division between the two. However, when the future becomes the present, the affective system kicks in, seeking immediate gratification and thereby seeking to override the rational system. In other words, the systematic deviations from rational decision making were understood as having resulted from a failed attempt by the individual’s rational system to control its affective system.

Within this general neuroeconomic approach behavioral economists have proposed different frameworks for the two systems. Let me give two prominent examples. In the paper that derived from his Nobel lecture [Kahneman (2003)], Kahneman argued for and employed the following framework.
Kahneman’s framework is an intriguing mix of psychophysics, neuroscience, and the desire to accommodate a distinction between two decision making processes. On the bottom row we see a distinction between two kinds of input for decision making: current external stimulation and information already present in the mind. These two sources of information form the input for two cognitive systems as described in the middle row: an intuitive system (also more neutrally labeled system 1), and a reasoning system (also more neutrally called system 2). The distinction between the two systems is mainly made in terms of the effort it costs to operate them. The reasoning system requires much effort and is relatively slow. The intuitive system operates much more quickly and is relatively effortless. Another distinction between the two is that between non-voluntary impressions in system 1, and voluntary judgments in system 2. The top row further distinguishes between perception and the intuitive information processing of system 1. The distinction between the two information processing systems and perception is made in terms of automaticity and accessibility. On the left we find decision making that is fully automatic and inaccessible. One example is the perception/decision regarding which of two rooms has a higher temperature, a perception/decision that is made automatically and without the individual having access to its process. At the other end of the spectrum, we find decisions that are made non-automatically and to which a large degree of
accessibility is possible. An example is the decision regarding which of two houses is preferred and hence will be bought. The intuitive system is a middle ground between these two and reflects decision making that often proceeds automatically, but which can also be accessed and altered, such as the decision between €3,000 for certain and a 0.8 chance at €4,000. Using this framework Kahneman thus brought together the automatic perceptual system investigated by psychophysics and behavioral economics’ use of neuroscience, while at the same time allowing for the possibility of deviations from rational economic decision making.

Another prominent framework, extensively discussed in what may safely be considered as one of the canonical articles in neuroeconomics, Camerer, Loewenstein, and Prelec (2005), “Neuroeconomics: How Neuroscience Can Inform Economics” is the following:

<table>
<thead>
<tr>
<th>Controlled Processes</th>
<th>Cognitive</th>
<th>Affective</th>
</tr>
</thead>
<tbody>
<tr>
<td>serial</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>effortful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>evoked deliberately</td>
<td></td>
<td></td>
</tr>
<tr>
<td>good introspective access</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Automatic Processes</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>parallel</td>
<td>III</td>
<td>IV</td>
</tr>
<tr>
<td>effortless</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reflexive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no introspective access</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Camerer, Loewenstein and Prelec (2005), p.16

In this representation the two systems are either controlled or automatic. Examples of automatic, cognitive processes, quadrant III, are judgments of relative temperature and shapes of objects. The automatic affective system, quadrant IV, depicts the pleasure and pain system that on the basis of the information provided by quadrant III and on the basis of information of past experiences attaches a value to the object. Quadrant I and II constitute the cognitive and affective part of the decision making process that can be controlled. A decision-maker may very much like to buy a car, but in quadrant I reason determines that he or she cannot afford it. And he or she may not
be hungry at all, while in quadrant II, not wanting to disappoint his or her friend’s cooking efforts, he or she takes a bite nevertheless.

Thus, in behavioral economics and its neuroeconomics research, descriptive real-world decision making should be seen as a struggle between a cognitive system that seeks a rational solution and an affective system that disregards the optimal decision in favor of immediate gratification. In that respect, the dual system approach of contemporary behavioral economics is different from the thinking of many other neuroscientists and social scientists who do not see the cognitive system as superior to the intuitive, emotional, or affective system. Examples of the latter include the work of Damasio [e.g. Damasio (2003)], research descending from Simon [e.g. Gigerenzer et al. (1999), Gigerenzer and Selten (2001)], and evolutionary theory inspired science [e.g. Barkow, Cosmides and Tooby (1992)].

The intertemporal choice and dual systems approach literature illustrates that in building behavioral economics, behavioral economists have incorporated the normative-descriptive distinction in their theories of economic decision behavior. They have come to see the normative not as something external to the individual, but as a rational system side by side an affective system, with which it strives for dominance. As Kahneman has also contributed to this new and changed perception of the normative and descriptive, it should be seen as reflecting a development in the behavioral decision research community at large, not only a development in behavioral economics.

2.2 Are preferences innate to human nature or do they emerge through interaction with the environment?

The incorporation of the distinction made between the rational norm and its imperfect realization in the economic agent considerably broadened the scope of behavioral economics. However, there seemed to be limits to broadening the scope as well. This can be illustrated by behavioral economists’ cooperation with anthropologists on the subject of the emergence of preference. This brief collaboration shows how behavioral economists, after their initial enthusiasm retreated when it turned out that this collaboration resulted in research that was at odds with the fundamental behavioral economic assumption of a fixed, universal benchmark of full rationality. As in the case of intertemporal choice, behavioral economists were unwilling to give up this benchmark of full rationality. The research was a large scale interdisciplinary
study of the ultimatum-game in fifteen small-scale societies. It was published in a number of journals. The most extensive discussion can be found in the book devoted to it, Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr and Herbert Gintis, *Foundations of Human Sociality* (2004). A reflection on the research summarized by Henrich can be found in Gigerenzer and Selten (2001).43

The motivation for this large interdisciplinary study was the following. The ultimatum game had been played all over the world, and has always led to the result that individuals do not play the rational optimum, but typically divide the money about half-half.44 However, the experiments had only been done with university students in advanced capitalist economies. The question was thus whether the results would hold up when tested in other environments.

The surprising result was not so much that the average proposed and accepted divisions in the small-scale societies differed from those of university students, but how they differed. Roughly, the average proposed and accepted divisions went from [80%,20%] to [40%,60%]. Individuals in different societies thus showed a remarkable difference in the division they proposed and accepted. Henrich and his fellow researchers correlated these differences with two economic characteristics of small-scale societies. First, they documented how much a group’s (normally the family) economic welfare depended on co-operation with other groups within a small-scale society. In this respect the societies differed greatly from almost none to almost completely. Second, the researchers investigated how much the group’s economic welfare depended on market exchange. There were also differences in the level of market integration. The researchers concluded that differences found in the behavior of individuals belonging to the different societies in the game should be attributed to differences in the environment in which they lived. As a consequence, preferences were not exogenous, but determined by the environment. Henrich and his collaborators stated this explicitly in a brief summary of their research in the *American Economic Review*:

---

43 The research was funded through the MacArthur Foundation’s research network *The Nature and Origin of Preference*. The network, “headed by Robert Boyd and Herbert Gintis received two grants: $2.55 million in 1997 and $1.8 million in 2002” [email Richards to author, December 4, 2008].

44 The ultimatum game is the following: Player one proposes a division of a fixed sum of money, player two either accepts (the money is divided according to the proposed division), or rejects (both players get nothing). If both players are rational player, one proposes the smallest amount possible and player two accepts.
[...] preferences over economic choices are not exogenous as the canonical model would have it, but rather are shaped by the economic and social interactions of everyday life. This result implies that judgments in welfare economics that assume exogenous preferences are questionable, as are predictions of the effects of changing economic policies and institutions that fail to take account of behavioral change. [Henrich et.al. (2001), p.77]

Giving up the exogeneity of preferences would have far reaching implications. It would mean that behavioral economists had to give up a fundamental tenet of economics, namely that preferences are given. That would have in turn far-reaching consequences for the theory of decision making itself. It would draw behavioral economists closer to cultural anthropologists, such as Boyd or Henrich, who conceive culture in evolutionary or ecological terms. Taking this approach and the results of the experiments seriously would imply not only that individual preferences to a large extent are determined by the environment and by learning, but it would also undermine the notion of fixed norms in behavioral decision research and well-defined rationality in economics. It is probably due to these extensive implications that in spite of plans to continue the research, follow-up studies and further elaboration of the implications of the experiments by this diverse group of economists and anthropologists to date have not been worked out.

2.3 Understanding behavioral economics in terms of rationality
When behavioral economics expanded, behavioral economists were both faithful to the Kahneman and Tversky legacy, while at the same time they sought to broaden its scope. Problematic in this regard were the labels of normative and descriptive, which were considered confusing in an economic context that already had created its own understanding of these concepts. We have seen examples of how economists changed this distinction by developing different models of economics decision making. As a consequence, behavioral economists in the 1980s-2000s have reinterpreted the normative-descriptive distinction in terms of rationality.

As set out in Chapter five, Thaler was well aware of the fact that the re-interpretation of economics in terms of normative versus descriptive raised the question concerning the definition of the descriptive theory when the normative theory is about rational behavior. And here Thaler was not very specific, or at least he
did not offer a conclusive answer. Thaler referred to behavior that deviates from the normative solution on a number of occasions as “irrational” or “non-rational.” Furthermore, he noted that he “would not want to call such choices rational” [Thaler (2000), p.138]. On other occasions Thaler referred to the normative-descriptive distinction as rational versus emotional [see e.g. Shefrin and Thaler (1988), p.611].

But the main interpretation Thaler used in the 1980s and 1990s was the term “quasi rationality,” most prominently as the title of a collection of articles, *Quasi Rational Economics* (1991). Quasi rationality suggests a category of behavior somewhere between the full rationality of the normative decision and irrational behavior. Regularly used in the 1980s and 1990s quasi rationality is perhaps best understood as the failed attempt of people to be rational, which is exemplified by the one suggested definition of the term that Thaler has provided: “quasi rational, meaning trying hard but subject to systematic error” [Thaler (2000), p.136]. On another occasion it was characterized as “less than fully rational” [Thaler (1991), p.xviii].

From the early 2000s onwards, the term which has been increasingly favored by behavioral economists is “bounded rationality.” The distinction that was made is that between the fully rational decision and the decision actually made that has been deemed boundedly rational when deviating from the rational decision. Full rationality in behavioral economics was defined as follows:

The standard approach in economics assumes “full rationality.” While disagreement exists as to what exactly full rationality encompasses, most economists would agree on the following basic components. First people have well-defined preferences (or goals) and make decisions to maximize those preferences. Second, those preferences accurately reflect (to the best of the person’s knowledge) the true costs and benefits of the available options. Third, in situations that involve uncertainty, people have well-formed beliefs about how uncertainty will resolve itself, and when new information becomes available, they update their beliefs using Bayes’s law – the presumed ability to update probabilistic assessments in light of new information. [Camerer et al. (2003), 12-14-1215]
Using the distinction between full and bounded rationality naturally entailed making references to the work of Herbert Simon. This can be illustrated by Gabaix and Laibson’s article “A Boundedly Rational Decision Algorithm” (2000). In this article, Gabaix and Laibson created decision algorithms for a specific decision problem that deviated from the “fully rational” algorithm, to which the term “boundedly rational” was applied. The following simulation was constructed. Consider a five-level decision tree where at each level, including the first, there are ten different pay-off boxes and between one and five connections to the next level. Each of the pay-off boxes’ possible connections to the next level has a certain probability, with the added probability of the connections of one box always amounting to one. An example of such a randomly generated decision tree appears as follows.

Source: Gabaix and Laibson (2000), p.434

Gabaix and Laibson argued that for a fully rational decision algorithm, the decision maker would be required to calculate the expected pay-off of each box and then make
a decision about which of the ten first level boxes to start. However, this would require a considerable amount of computational capacities and time. The authors accepted from Simon’s work that both are often unavailable. They considered three “boundedly rational” alternative algorithms: i) “Follow-the-leaders” (FTL), which ignores probability paths of less than 0.25, ii) column cut-off, which ignores one or more of the last columns, and iii) discounting, which discounts the values of later columns. Of these three alternatives, FTL came closest to the fully rational outcome. Furthermore, it closely matched behavior of subjects in an experiment who were faced with the same task. It was concluded that individuals’ behavior might be explained in terms of a boundedly rational algorithm.

In other words, the concept of bounded rationality was taken from Simon and together with the concept of full rationality employed to rephrase Kahneman and Tversky’s normative – descriptive distinction. In one clear sweep, Kahneman and Tversky’s distinction between the concepts of normative-descriptive were replaced by concepts more appropriate in an economic context, and at the same time Simon was appropriated as an authoritative source for the use of these concepts. Simon provided behavioral economics with the fitting language and offered it the necessary authority. This is exemplified by Kahneman in the American Economic Review article based on his Nobel lecture, entitled “Maps of Bounded Rationality: Psychology for Behavioral Economics” (2003). In the opening passage Kahneman used the term “bounded rationality” and referred to Simon, but at the same time subtly but clearly distinguished his and Tversky’s work from Simon’s.

The work cited by the Nobel committee was done jointly with Amos Tversky (1937-1996) during a long and unusually close collaboration. Together, we explored the psychology of intuitive beliefs and choices and examined their bounded rationality. Herbert A. Simon (1955, 1979) had proposed much earlier that decision makers should be viewed as boundedly rational, and had offered a model in which utility maximization was replaced by satisficing. Our research [on the contrary – FH] attempted to obtain a map of bounded rationality, by exploring the systematic biases that separate the beliefs that people have and the choices they make from the optimal beliefs and choices assumed in rational-agent models. [Kahneman (2003), p.1449]
In a clever way, Kahneman invoked Simon to construct authority for the behavioral economic program, while at the same time interpreting the concept of bounded rationality in such a way that it would become fully compatible with his and Tversky’s approach and that of the behavioral economists. Thus, at the same time behavioral economists remained committed to Kahneman and Tversky’s methodological distinction between separate normative and descriptive domain, re-labeling this distinction in more appropriate economics terms, and claiming Simon’s intellectual heritage as adding credibility to their research.

2.4 Economic policy as a form of paternalism

The re-interpretation of Kahneman and Tversky’s distinction between the normative and the descriptive in terms of a conflict within the economic decision maker had important consequences for welfare economics. Most mainstream economists in the 1990s and 2000s associated welfare economics in one way or another with the term normative. That was one reason why Kahneman and Tversky’s labels of normative and descriptive invoked confusion when inserted into economics discourse. The reinterpretation of normative versus descriptive in terms of full-rationality versus bounded rationality solved this confusion and in turn allowed behavioral economists to develop their own position on welfare economics. Behavioral economists described their new stance on welfare issues as a form of paternalism. Behavioral economic paternalism, then, resulted from the solution that behavioral economists found in dealing with the violations of the normative or full-rationality benchmark. It illustrates the ultimate consequence behavioral economists have drawn from their new program.

The paternalism debate involves the application of behavioral economic insights into real-world policy questions. For instance, behavioral economics had discovered that people often save much less for their pensions than they should, and that when they do save, they do not diversify their portfolios optimally. Following on these results, programs have been set up to investigate how people can be induced to save more for retirement and better diversify their stock portfolios [e.g. Cronqvist and Thaler (2004), Thaler and Benartzi (2004)].

Another example concerns the use of medication. It has often been found that people who need to take drugs on a regular basis are very lax at doing so. Even when the risks are substantial and potential costs in terms of health very great, such as in the case of medication that reduces the chance of having a second stroke, people are very
lax at taking their medication properly. To solve this problem, programs have been set up that investigate how insights from behavioral economics can be used to design incentive mechanisms that induce people to take their medication [e.g. Badger et al. (2007)]. Finally, behavioral economists have turned their attention to development economics, with the purpose of using insights from behavioral economics to improve the functioning of development programs [see e.g. Mullainathan (forthcoming), and Betrand, Mullainathan and Shafir (forthcoming)]

Behavioral economists have framed and defended this research in a number of closely related ways. Well-known is Thaler and Sunstein’s (2003) “Libertarian Paternalism.” Libertarian paternalism can be understood as a paternalism that does not restrict individual freedom of choice. Thaler and Sunstein distinguish themselves explicitly from the Samuelsonian stance towards welfare issues.

We clearly do not always equate revealed preference with welfare. That is, we emphasize the possibility that in some cases individuals make inferior choices, choices that they would change if they had complete information, unlimited cognitive abilities, and no lack of willpower. [Thaler and Sunstein (2003), p.175]

In the behavioral economics paternalism debate, the justification for paternalistic policies has been the fact that decisions people actually make, their “revealed preferences,” do not always match with their “true” preferences. Behavioral economists have thus constructed a distinction between “revealed” and “true” preferences. This does not mean that preferences are context dependent. Rather, it means that it depends on the context whether the true preferences can and will be revealed appropriately. A source that is sometimes relied on in this regard is John C. Harsanyi who noted that “in deciding what is good and what is bad for an individual, the ultimate criterion can only be his own wants and his own preferences,” where the individual’s “own preferences” are his “true” preferences: “the preferences he would have if he had all the relevant factual information, always reasoned with the greatest possible care, and was in a state of mind most conducive to rational choice” [quoted in Angner and Loewenstein (forthcoming), pp.53-54].

A more detailed and elaborate explication and defense of this new branch of behavioral economics can be found in Camerer et al. (2003) “Regulation for
Conservatives: Behavioral Economics and the Case for ‘Asymmetric Paternalism’.” In this article, the five authors (Camerer, Issacharoff, Loewenstein, O’Donoghue and Rabin) make a case for what they label “asymmetric paternalism,” where “[a] regulation is asymmetrically paternalistic if it creates large benefits for those who make errors, while imposing little or no harm on those who are fully rational” [Camerer et al. (2003), p.1212]. Behavioral economics, then, “describes ways people sometimes fail to behave in their own best interests” [Camerer et al. (2003), p.1217]. These “apparent violations of rationality [...] can justify the need for paternalistic policies to help people make better decisions and come closer to behaving in their own best interests” [Camerer et al. (2003), p.1218].

Thaler and Sunstein (2003) countered possible aversions to paternalism by economists and others by linking paternalism to libertarianism. Camerer et al. (2003), on the other hand, founded their defense of paternalistic policies on the need for asymmetry in paternalistic policy. The definition of asymmetric paternalism resembles the Paretean improvement argument: “a policy is asymmetrically paternalistic if it creates large benefits for those people who are boundedly rational [...] while imposing little or no harm on those who are fully rational [Camerer et al. (2003), p.1219, emphasis in the original]. Or, in other words, “asymmetric paternalism helps those whose rationality is bounded from making a costly mistake and harms more rational folks very little” [Camerer et al. (2003), p.1254]. Another way of putting it, Camerer et al. (2003) argued, is to see the limitedly rational individual as imposing negative externalities on his or her own demand curve. “When consumers make errors, it is as if they are imposing externalities on themselves because the decisions they make as reflected by their demand do not accurately reflect the benefits they derive” [Camerer et al. (2003), p.1221]. Hence, there is a need for a policy maker who can remove the externalities and redirect behavior in such a way that the externalities disappear. Camerer et al. (2003) furthermore noted that firms could either consciously or unconsciously use the irrationality of individuals to gain more profit.

A detailed example has been provided by Grubb (2007), who showed that cellphone companies could permanently increase their profits using the phenomenon of overconfidence. When consumers systematically underestimate the number of minutes they will use their cell phones each month, it is profitable for firms to charge the marginal costs per minute up to the number consumers expect they will use their
cell-phones. After this point, they can greatly increase their rates. Consumers will not mind because they do not expect to be using their phones that much. But the cell-phone company knows that the consumers are overconfident, and thus knows the customers will use some of these expensive minutes. Thus the cell-phone company increases its profit by exploiting consumers’ overconfidence. The example shows that firms in the market can use the limited rationality of individuals and thus not produce efficient outcomes. Moreover, the example shows that it might be more in the interest of companies to maintain or even amplify the limited rationality of individuals. It suggests that firms are apt to look for ways to increase the irrationality of individuals, and that as a result the market, instead of producing an efficient equilibrium, could even produce greater irrationality. Behavioral economists have shown that deviations from full rationality do not only persist in markets, but that markets can even increase this deviating behavior.

On the basis of these results, behavioral economists have argued that economists are morally obliged to act against the violations of full rationality:

As economists, how should we respond to the seemingly self-destructive side of human behavior? We can deny it, and assume as an axiom of faith that people can be relied upon to do what’s best for themselves. We can assume that families paying an average of $1,000 per year financing credit card debt are making a rational tradeoff of present and future utility, that liquidity constraints prevent investing in employer-matched 401k plans, that employees prefer investing in their own company’s stock instead of a diversified portfolio [...] that people are obese because they have calculated that the pleasures from the extra food, or the pain of the foregone exercise, is sufficient to compensate for the negative consequences of obesity. [Loewenstein and Haisley (forthcoming), p.4]

According to behavioral economists, economics is particularly suited for solving the violations of full rationality because it possesses the knowledge of how to “steer human behavior in more beneficial directions while minimizing coercion, maximizing individual autonomy, and maximizing autonomy to the greatest extent possible” [Loewenstein and Haisley (forthcoming), p.6]. The role of the economist in this regard can be seen as analogous to the psychoanalytical therapist. “Just as the
therapist endeavors to correct for cognitive and emotional disturbances that detract from the well-being of the patient, such as anxiety, depression, or psychosis, the economist/therapist endeavors to counteract cognitive and emotional barriers to the pursuit of genuine self-interest” [Loewenstein and Haisley (forthcoming), pp.9-10].

Behavioral economists have attempted to solve mankind’s limited rationality problem by using phenomena similar to those that formed the basis for behavioral economics to begin with. The reason that they could do so was that behavioral economics had remained faithful to Kahneman and Tversky’s approach. The most important phenomenon in this regard is what is most commonly known in behavioral economics as framing. One of the central findings of Kahneman and Tversky’s behavioral decision research and behavioral economics was that people are susceptible to the way in which a choice is presented to them. Depending on the “reference point,” in Kahneman and Tversky’s terms, or “frame,” the term Thaler favored for behavioral economics, people change their preferences. The example taken from Thaler and Sunstein (2003) is of the cafeteria manager who can either place the desserts before the fruits or vice versa. If he or she frames this decision as fruits-before-desserts, the fruit will be chosen more often. Thus, framing is used to influence people’s behavior without affecting their freedom to choose in any significant way. Changing the default option from not-participating to participating in pension saving schemes is another often-quoted example.

By exploring how policies can be designed to solve the bounded rationality of individuals, behavioral economists have taken the full-rationality versus bounded rationality framework and the experimental results of psychologists and economists to their ultimate consequences. Behavioral economic paternalism is very an economics solution to bounded rationality, emphasizing incentive mechanisms and monetary rewards.

3. Defining behavioral economics
The different explorations of behavioral economists in the 1990s and 2000s contributed to the gradual construction of a clearly defined field. Over the course of some twenty years, behavioral economics used the conceptual core as laid out by Kahneman and Tversky and explored its implications in rethinking the economic agent. Rather than the received rational decision maker of neoclassical theory, the individual became a decision maker in whom the cognitive and the affective system
were in competition. Despite the fact that behavioral economists questioned the
normative benchmark at different points in time, in the end this distinction remained a
core element of behavioral economics, though in a modified form. Over time the
labels were re-interpreted in terms of full and bounded rationality. In turn, this
allowed behavioral economists to claim the heritage of Simon, though they took a
different perspective than other social scientists who claimed to work in the tradition
of Simon, namely what could be broadly and roughly labeled the ecological approach
to rationality. This group includes anthropologists such as Boyd and Henrich,
psychologists such as Gigerenzer, and experimental economists such as Vernon Smith
and Charles Plott. But behavioral economists’ use of Simon also made them more
clearly aware of how they differed from other strands involved with decision research.

These theoretical and conceptual developments formed only one aspect of a
gradual process which defined behavioral economics. In addition to this theoretical
and conceptual definition, behavioral economists came to define themselves across
disciplinary lines. In particular, they gradually began to distinguish behavioral
economics from the discipline of psychology and from the sub-discipline of
experimental economics. In the case of psychology, the reason was that behavioral
economists wanted to be accepted by the economic mainstream. To achieve this they
had to comply with the mainstream economic view that economics and psychology
are separate disciplines and that economics is superior to psychology when it comes to
rigor and formal modeling. Some aspects regarding the attempts made by behavioral
economists to distinguish themselves from psychology have been hinted at above in
the discussion concerning Laibson’s way of incorporating psychological insights.
More specifically, behavioral economists have employed the following arguments to
distinguish behavioral economics from the discipline of psychology.

First, behavioral economics has been defined as economics on the basis of its
use of mathematical modeling. This argument has remained largely implicit, as in the
case of Laibson, but it is nevertheless clear that behavioral economists consider their
use of mathematics superior to that of the psychologists, and that they see it as a
defining characteristic of economics. Another illustrative example in this regard is the
work of Matthew Rabin [e.g. Rabin (1993, 1994, 1998)]. Rabin has incorporated
experimental results from psychology into economics, but has combined this with
advanced economic mathematical modeling. He argues that “none of the broad-stroke
arguments for inattention to psychological research are compelling,” but at the same
time cautions that “not all psychological research will be [...] proven to be of great economic importance.” The reason is that because they come from psychology, these psychological results have not yet been subjected to “the same rigorous standards that our discipline, at its best, applies elsewhere” [Rabin (1998), p.41]. That is, ultimately it requires the mathematics of economics to judge how useful the insights from psychology are. This use of mathematics is something that defines behavioral economics as economics, and therefore as different from psychology.

Second, behavioral economists distinguished themselves from psychology on the basis of their use of the experimental method. For instance, in a methodological comment on sharing data or experiment instructions with other researchers, Camerer noted that “[i]f you asked a psychologist for data or instructions he or she might be insulted, because the convention in that field is to give the writer the benefit of the doubt” [Camerer (2003), pp.34-35].

Another example concerns the use of deception. A common experimental procedure in psychology is to tell the experimental subjects the experiment is about one thing, while it is in fact about something else. Behavioral economists and experimental economists have resisted this method of deception ever since experiments have been used in economics. The reason for instance that Grether and Plott (1979) controlled for the fact that the experimenters had been psychologists was that they suspected that because the subjects knew they might be misled in a psychological experiment they would behave differently. In recent years the use of deception has produced increasing discussion in the experimental communities of economics and psychology [e.g. Hertwig and Ortmann (2008), Ortmann and Hertwig (2002), Jamisona, Karlaub, and Schechter (2008)].

The motivation for explicitly drawing a border between behavioral and experimental economics was different. Both research programs were part of the same economic science, and to define behavioral economics as economics, it was not necessary to distinguish between behavioral and experimental economics. The reason that behavioral economists nevertheless wanted to distinguish their field explicitly was that they had come to realize that theoretically they were of a fundamentally different opinion than experimental economists regarding the relation of individual decision making to market outcomes. However, because experimental economists used the same experimental method as behavioral economists, outsiders could easily

45 Interesting in this regard is that Camerer holds a PhD (defended in 1981) in behavioral decision research, and not in economics. His supervisor at the University of Chicago was Robyn Hogarth.
conclude that they must be the same. Therefore, the two needed to be clearly
distinguished. The distinction was drawn on the basis of the experimental method.

The different views of experimental economics by behavioral economists have
been grouped together in Loewenstein (1999) “Experimental Economics from the
vantage-point of Behavioural Economics.” Loewenstein (1999) positioned behavioral
economics explicitly in opposition to experimental economics. He formulated his
critique in terms of the “psychological distinction” of external versus internal validity.
Under the heading of external validity, Loewenstein saw four problems with
experimental economics. First, experimental economics put great emphasis on the use
of auctions in its experiments. As people in reality hardly ever find themselves in an
auction situation, it is doubtful that these experiments can tell us very much about
economic behavior in the real world. Second, Loewenstein disagreed with
experimental economists’ use of repetition in what has been called the Ground Hog
Day argument. In reality, Loewenstein argued, people never make the exact same
decision forty times in a row. Real world behavior is much more like the first few
rounds of an experiment than the last two or three rounds. Third, Loewenstein
criticized experimental economists for their tendency to reduce real-world content to
the absolute minimum possible. Apart from the fact that a context-free experiment is
an illusion, Loewenstein argued it also greatly reduces the external validity of the
experiments.46 Instead, economists should, just as Loewenstein himself, make the
experimental situation as congruent with reality as possible; hence make the
experiment “context-rich.” Fourth, according to Loewenstein experimental
economists wrongly assumed that monetary rewards result in strict control over
incentives. It has been shown in numerous experiments that this is not the case. With
monetary incentives, subjects are also likely to be driven by other motives than profit
maximization. Finally, one problem concerning internal validity that Loewenstein
observed was that experimental economists had been far too careless in not using
randomization and in comparing the experimental results that had been obtained under
different circumstances.

---
46 Loewenstein both uses ‘context’ and ‘content.’
4. Behavioral economics in the 2000s: stable, defined, but slightly schizophrenic

In the 1990s and 2000s behavioral economics was built, defined, and distinguished as a stable mainstream economic program. It was built on and remained true to the framework as laid out by Kahneman and Tversky, although it re-labeled its key concepts in terms of rationality. On the basis of their explorations, behavioral economists came to define step by step behavioral economics both theoretically and conceptually as an economic program that seeks to develop theories of bounded rationality in order to complement the normative benchmark of full rationality. In turn, this has laid the foundation for a new approach to welfare policy under the rubric of paternalism.

But the emerging behavioral economics has also been somewhat schizophrenic. Behavioral economists have defined behavioral economics as an important improvement on the traditional neoclassical economic paradigm by this incorporation of insights from psychology, but at the same time they have sought to define behavioral economics as being different from psychology. They have claimed the heritage of Simon and his bounded rationality, but at the same time they have explicitly distinguished behavioral economics from the work of other social scientists using Simon’s work. Yet, they have not done so by discussing how Simon should be interpreted, but instead on the basis of their view of the proper way to conduct experiments in economics. And finally, they have sought to define behavioral economics as mainstream economics, but at the same time they have gone to great lengths to embed behavioral economics more firmly in the cognitive and neurosciences.

The reason for this slight schizophrenia is that the incorporation of psychology has not been a neutral process. In fact, behavioral economists have used psychology to redefine economics. The psychological theories, experimental results and authoritative figures have not been used for their own sake but solely because they could be used to steer economics in a different direction. Part of this new direction has been a language of “enriching” economics with psychology, but this is merely a rhetorical guise for convincing economists which part of psychology they should understand, and in which way that should alter economic reasoning. That behavioral economists have not been neutral receivers of psychological input, but have actively built, defined and distinguished behavioral economics using psychology does not
diminish or undermine the behavioral economic project. However, it does raise the question regarding how we ought to understand the flow of theories, methods and results that cross from one discipline to the other.