Analyse & Kritik 29/2007 (© Lucius & Lucius, Stuttgart) p. 60–85

Bernhard Neumärker

Neuroeconomics and the Economic Logic of Behavior

Abstract: Recent neuroeconomic studies challenge the conventional economic logic of behavior. After an introduction to some starting points of brain research in 'classical' economics we discuss the final and contingent causes of rational and irrational behavior in neuroeconomics and standard economics and present the concept of expanded rationality models (ERM) which imports neuroeconomic elements like emotions, beliefs and neuroscientific constraints and exports improved testable predictions. The typical structure of neuroeconomic proof of economic models and the imprecise neuroscientific measurement let us suggest a feed-back structure of economic research. We apply Hirshleifer's conception of macro-/micro-technology and contrast it to the neuroeconomic black box critique on economic theory. Furthermore, instead of direct adjustment of preference structures we propose the auxiliary creation of neuroeconomic constraints like action-dependent or outcome-dependent neuroeconomic belief constraints (NBC) and emotion compatibility constraints (ECC). This prepares the ground for our examination of neuroeconomics and ERM as two different paths towards an analytical unification of behavioral sciences.

0. Introduction

Standard economic theory and its application study individual and collective decision-making but not the cerebral processes of individual decisions and human behavior. Neuroscientists note that an understanding of brain procedures is essential for applied behavioral theory and real-world research on human behavior. Especially, the real-life human registration and processing of terms like "utility", "reward" and "cost" is an apparently unsolved question. Neuroeconomic research indicates departures from the well-known essential of economic theory, namely the model of the rational self-interested actor. Early economists saw that scientific obligation to discuss and contingently integrate more realistic behavioral pattern in their research agenda. Smith's *The Theory of Moral Sentiments* (1759) rested upon the emotion of sympathy. In his *Mathematical Psychics*, Edgeworth (1881, 104) accepted that "man is for the most part an impure egoist", and Hayek, the Nobel laureate always being interested in constitutional conditions of individual, political and social life, published with *The Sensory Order* (1952) an investigation into the nature of human mental events, the role of the nervous system in determining principles and systems of order inside the brain and the consequences of these human factors for the social order.

The presentation of Edgeworth on the list seems to be surprising because Edgeworth affirmed that selfishness should be at first the only relevant motive applied in economic theory and labeled this notion "the first principle of economics" (Edgeworth (1881, 16). And relying on the first quotation from Edgeworth's work, Sen (1977, 317) asked "why Edgeworth spent so much of his time and talent in developing a line of inquiry the first principle of which he believed to be false". By contrast, Colander (2007, 216–219) reviews Edgeworth's conception of the physio-psychological hedonimeter that would not only measure utility but also integrate egotistic and altruistic happiness of the person. Edgeworth was certainly aware of the limits of his first principle but hedonic measurement was not feasible. Therefore, he and his followers focused economic research on the implications of the first principle.

Hayek's (1952) subtitle *An Inquiry into the Foundations of Theoretical Psychology* reports clearly Hayek's opinion on the kind of subject. In a review of his own contribution (Hayek 1982), he stated clearly that *The Sensory Order* was inspired by his interest in psychological conditions of human life, and that research on *The Sensory Order* was never intended to be highly related to economic theory. Smith, on the other hand, separated his research in a more formal way the first part of which was focused on emotions, passions and sentiments whereas the second part published in *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776) presented the politico-economic logic of human behavior and social life. Smith, Edgeworth and Hayek emphasized at least implicitly some duality between economics and psychology[1] but it seems that presumably two of them, namely Smith and Edgeworth, got the creation of an integrated and presumably unified approach of human behavior straight. And they all had in common that for modifications of rational actor models psychological factors would play an enormous role.

Emotions and mental events like sympathy actuating benevolent behavior, impure egoism or some kind of altruism are central subjects for neuroeconomic inspection. The concept of social preferences widely accepted in behavioral economics offers the starting point for reconciliation between pure self-interest and other-regarding preferences. And neuroscientific measurement of the activity of reward- and fear-related brain areas open, in principle, the door for a new technique of utility measurement. This provides a basis for the optimistic adoption of Smith and Edgeworth in neuroeconomic papers. "One may wonder whether Adam Smith, were he working today, would not be a neuroeconomist" (Rusticchini 2005b, 205), and "Edgeworth would have been a strong supporter of neuroeconomics work with brain scans to relate experience to a person's pleasure and pain" (Colander 2007, 224).

The sensory theories of Hayek and Smith are based on the assumption that introspection is the only way of analyzing mental events. Very limited technical support of psychological and neural inspection of the human body at the days of Smith and Hayek restricted their scientific possibilities, and Hayek

[1] Hayek (1952, 1–8, 191–193; 1982, 291) discussed the duality between the physical and the psychological order and Edgeworth based utility measurement on experimental studies in psychophysics (cf. Colander 2007, 217).

(1952, 191–194; 1982, 292) forcefully criticized the logic of introspective research. He concluded, that "the whole idea of the mind explaining itself is a logical contradiction—nonsense in the literal meaning of the word—and a result of the prejudice that we must be able to deal with mental events in the same manner as we deal with physical events" (Hayek 1952, 192). Hayek's pessimistic view of explaining the mind contradicts clearly Edgeworth's physical measurement of emotions as Hayek claimed that theoretical psychology could never "enable us to substitute for the description of particular mental events descriptions in terms of particular physical events" (Hayek 1952, 191). From this essential proposition it follows that we shall never manage a perfect unification of behavioural sciences.

Whether neuroeconomic research design contributes to the solution of Hayek's problems of logical contradiction and practical dualism between physics and psychology is a task for future research. It is still an open question, if the final result of neuroeconomic research is a unification of behavioral theories or a verification of multiple professions with dual- or multiple-process approaches of the brain. In section 3 we present the neuroeconomic perspective on a unification of behavioural sciences. But, at the very least, neuroeconomic analysis rediscovers and enhances the research agenda of early psycho/socio-economic studies in human behavior. That is, the neuroeconomic approach is deeply rooted in former research of three important founding fathers of comprehensive economic analysis and can be viewed as a modern scientific attempt to revitalize the true human nature of economics and add behavioral components to mainstream economics.

In the next sections, we will not discuss the neuroscientific techniques and methods used in neuroeconomic research (see, for that, the overviews of Glimcher/Rustichini 2004; Camerer/Loewenstein/Prelec 2005, 11–14; Kenning/Plassmann, 2005) or the results in special fields like neuromarketing (cf. Kenning/Plassmann/Ahlert 2007) but rather the basics and the opportunities of neuroeconomic reasoning. Section 1 describes and critically reviews the differences between traditional economics and neuroeconomics. In section 2, we examine the way neuroeconomic data could improve economic theory. The next section (3) deals with the potential of neuroscientific integration and unification of different behavioral sciences and section 4 concludes.

1. The Causes of Human Behavior in Economics and Neuroscience

The human brain is the natural center and mechanism of individual decision-making. Economists apply a highly stylized model of decision-making procedures inside the brain in that they usually assume that individuals are rational and will act in their own interest. The basic idea of rational choice is the optimization of a preference function defining individual goals subject to different beliefs and opportunities.[2] The economic man wants to achieve the highest (expected) net benefit. Other decisions are called irrational and outside the scope of economics.

[2] See, for example, material, non-material and informational incentive and feasibility constraints as well as optimistic or pessimistic attitudes or perceptions (cf. Hirshleifer 1998).

Even bounded rationality is based on the rational application of capacity constraints concerning information processing and environmental complexity. It is still a part of economic rationality (cf. Gintis 2006).

Behavioral economics criticizes this reading of the brain and integrates psychological and biological insights of human behavior. Neuroeconomics is a section of behavioral economics and studies neurobiological mechanisms and cerebral activity of economic decision-making and human behavior (Zak 2004, 1737). Modern neuroscientific techniques of functional neuroimaging like functional magnetic resonance imaging (fMRI) and positron emission tomography (PET)[3] are used to monitor certain brain responses to different tasks of economic decision-making. With these methods we know the brain areas that are active when some behavior is observed. Particular regions in the brain are of special interest for transdisciplinary research between neuroscientists and economists. The area of reward may report monetary and non-monetary gross or net benefits and the area of fear feelings of risk, expected losses and aversions like loss aversion or some inequality aversion.[4] Furthermore, it is possible to identify regions and cerebral mechanisms which are related to individual judgements under risk taking and payoff-discounting (Montague/Berns 2002; Sanfey at al. 2006).

Many results of neuroeconomic studies challenge the line of arguments in conventional economic analysis. Deviations from the assumptions, mechanisms and conclusions of pure economic theories seem to be many and various. One fundamental result of fMRI and PET tests is that individual 'decisions' and reactions concerning economic problems are heavily based on emotions (cf., instead of many others, Rustichini 2005a and Shiv/Loewenstein/Bechara 2005).[5] And if emotions can not be economically rationalized or parameterized, there is little reason to apply the paradigm of the optimization of pure self interest. Emotional sets drive an individual to act beyond the bounds of 'rational' self-interest. If emotions force actions in the interest of no one, then the assumption of rational behavior may be rejected.[6] And if emotions lead to malevolent or benevolent behavior, then the assumption of human selfishness is of limited use.

First important criticism that is raised to the method and the behavioral foundations of standard economic theory was developed in economics itself by Sen (1977) and his rational fools' case. And it has long been appreciated in modern economic theory that emotions, passions and feelings can be introduced as a part of preferences (Becker 1976; Hirshleifer 1987; Andreoni 1989; Gintis 2004a,b), the key point being that individuals can be, e.g., emotionally malev-

---

[3] An older technique is electromagnetic recording (Kenning/Plassman 2005, 344).

[4] Negative emotional processing may have different causes. E.g., a person dislikes advantageous income distributions by showing the emotion of guilt. Or, fear may be associated with risk aversion.

[5] The other fundamental non-economic brain mechanism is automatic processing facing controlled behavior (cf. Camerer/Loewenstein/Prelec 2005, 15–18). For our purposes, the study of emotional processes and affective drives is sufficient, because affective brain responses influence negatively cognition as well as control (cf. the definition of emotions in Sanfey et al 2006, 112).

[6] We can not identify some kind or direction of optimization. Even the fictive benevolent dictator with no self-interest optimizes 'for the public benefit', that is, in the weighted interests of other citizens.

olent or benevolent. Spite, malice, envy and hostility on the one hand and benevolence, kindness and amiability on the other hand are, then, specific factors of preferences determining individual utility functions. Other non-material motives and affective drives like prestige, vanity, anger, warm glow or honor can be important emotional factors as well (cf., for the case of prestige: Harbaugh 1998). Such extended economic models opening the research agenda of sociobiology (Becker 1976) or bioeconomics (Hirshleifer 1998; Gintis 2004b) simply expand the content of the preference function and assume rationality in the way that persons with emotions and passions still optimize in accordance to their goals without seeking a merely egotistic goal. This broadened concept of rationality implies at least other- as well as self-regarding behavior, but individual decisions are still rigorously correlated with reason. Actions in the short run or long run interest of no one are not incorporated in these 'expanded rationality models' (ERM).

From the strict perspective of neuroscience, emotions remove calculated optimization and imply some loss of (economic) control.[7] Emotions are no direct part of preferences but confuse preference-oriented behavior and decisions. They lead to irrationality in the sense of non-optimizing decision-making. Many neurobiological studies based on the modern technique of neuroimaging reveal that apparently irrational decisions are correlated with the emotional part of the brain (see, for example, Shiv/Loewenstein/Bechara 2005). This neuroeconomic view would end in a rejection of both rationality and selfishness. Automatic brain responses governed by emotions could serve the self interest and rationality of a person only by accident.

A mild reaction to that failure of the traditional economic approach is the integration of neurobiological features in economic theories of behavior as essential parameters of preferences and/or constraints. Emotions could be a part of intangible and intrinsic motives, impulses and goals. If the main task of the economist is to establish useful predictions on which human behavior will follow in different situations of individual and social life, these predictions will be improved, when the economist knows the set and the quality of specific neuroscientific parameters of behaviour in certain situations. The intermediate decision-making process of the brain links economic models of decision-making and strategic interaction with actual behaviour and economic outcomes and, thereby, improves the knowledge about the relevant parameters of human behaviour in isolated and social situations. This is of special interest for the formation of a theory on completely new economic problems. Neuroeconomic insights can improve the development and choice of relevant ERM.

One important attempt of this moderate neuroeconomic perspective is to identify emotions correlated with economic terms like "benefit", "cost" and "utility". The detected brain regions of reward and fear are first candidates for the measurement of 'economic' emotions. Neuroeconomics studies, then, not only the boundaries of rational economic behavior but also the identification and specification of the human rational actor. De Quervain et al. (2004), for ex-

[7] And this is also the regular starting point of sociological and psychological explorations into economic problems.

ample, detected preferences for punishing norm violations, because punishment activated reward-related brain regions even in situations of costly punishment where benefits would have to be weighed against the costs of punishing. People seem to derive non-monetary benefits from the act of punishment.[8] This shows the high potential of the neuroeconomic approach in detecting intangible benefits and costs. Thus it appears that neuroscientific measurement of emotions and other intangibles probably also play a crucial role for cost-benefit analysis. A vital problem of cost-benefit research is the identification and quantification of intangibles and intangible goods (Boadway/Wildasin, 1984; 201, 313). Nelson (1970) and Darby/Karni (1972) developed experience and credence dimensions of intangibles which are hard to evaluate. Emotions lead also to a non-material component in the valuation of public projects. For example, emotional effects originated from pleasure grounds influence clearly individual and social appraisal of a park. The value will be measured more accurately, if we can identify different degrees of brain activity in the reward area as a reaction on an announcement of different park improvements. Social preferences are also intangibles and may change the valuation of many public projects (Bowles 2007). Neuroeconomic studies of the valuation of public projects will, therefore, support the appraisal of the benefits and costs of public projects. But so far the neuroeconomic approach is lost more in the direct explanation of human behavior than in value determination (but see Montague/King-Casas/Cohen, 2006; Sanfey et al. 2006).

ERM and neuroscientific approaches both have their merits and limits. The first one expands the rational actor model whereas the latter adds important, sometimes irrational or seemingly irrational aspects of human behavior to the rational economic man. On the one hand, even expanded economic rationality can not explain or predict all facets of human behavior. On the other hand, the economic rational actor model faced with apparent deficiencies show a very high potential of its recovery, and the well-known 'as if' assumption of economic modeling (criticized by Rustichini 2005b; Camerer/Loewenstein/Prelec 2005, 10; Fehr/Fischbacher/Koesfeld 2005) allows new and alternative explanations of human behavior which may serve again as a basis for additional neuroscientific inspection. Neuroeconomists stress the fact that only experimental research can inspect effects of emotions as final causes of behavior. And that only neuroscientific measurement techniques prove and control the production of emotions and uncover various 'natural' mechanisms behind individual decisions. Economists see final causes as sources and factors of individual and social preferences. Conventional economists regard these factors as exogenously given and out of the scope of economics. They insist on the 'as if' approach and alter the sets of preference functions, beliefs and constraints. Other economists are interested in the set or a subset of the final causes. In bioeconomic studies, reproductive success is a crucial final cause of behavior and leads to the selection of those preferences that support "survival in a world of scarcity and competition" (Hirshleifer 1998, 457, cf. also: Camerer 2007, C31). Affective and visceral drives

[8] The intangible benefit of the punisher can be based on self-interest (e.g., sadism, glory, prestige) or on a positive social preference (benefits on third persons). We will discuss this neuroscientific problem of the specification of non-material benefits in a later chapter.

like hunger and sexual feelings promote reproductive success directly. Indirect non-monetary factors are prestige, dominance and honor. Reproductive success seems to be a genetic disposition but, as far as we know, it is not proved, if it is a genetic code for brain processes.

Neuroeconomic research is based on another subset of final causes, namely the emotions of reward and fear. They often lead to the selection of those preferences that support cooperation in a world of scarcity and competition.[9] These preference functions could contain benevolence (e.g., altruism or warm glow) or malevolence (e.g., violence or sadistic punishment). Preferences, beliefs and constraints are, then, contingent causes that establish actual behavior.

The conceptually unsolved problem for ERM is the assimilation of the neuroscientific protocol of irrational behavior. Actions in the interest of no one and, therefore, irrationality cannot be easily explained inside standard economics.[10] An important aspect of irrationality is the carrying out of behavior that we know is reducing (net) benefits. This leads constantly to a deviation of factual behavior from preference-related optimal choice. Furthermore, non-maximized preferences can be stated in cases of healthy individuals, addicts and persons with brain damages. Shiv/Loewenstein/Bechara (2005) exhibit that persons with prefrontal cortex damage participate much more in repeated lotteries with an expected profit than healthy persons. This means in specific situations of decision-making that individuals without brain damage may tend to be irrational. On the other hand, Smith/Tasnádi (2003) present a neuroscientific foundation for an economic theory of rational addiction so that addiction not necessarily challenges the logic of ERM.

Neuroeconomic experiments may indicate rational or irrational behavior. The key point mentioned above is that ERM can't explain anything and should not be used to rationalize obvious unreasonable behavior by adjusting preference functions or beliefs 'in the right way' (cf. also: Pesendorfer 2006, 720).[11] Neuroeconomic experiments may help to determine relevant preference functions so that economists can execute a better cross-check of the grade of rationality in behavior and choice and the accuracy of specific ERM. Due to the fact that emotions play the central role in neuroeconomics, the extension of preference functions toward other-regarding preferences is an obvious direction. But the additional or probably conflicting incorporation of other intangibles as pride, vanity and prestige seems to cover the notion of emotions in a better way.

2. Neuroeconomic Reasoning and Economic Logic

In this section, we discuss the ways neuroceonomics contributes to the development and refinement of economic thinking. The neuroscientific determination

[9] In addition, reward and fear specify time preferences (hyperbolic discounting) as well as preferences for risk taking (nonlinear probability weighting).

[10] Small irrationality is applied as an exogenous factor in game-theoretic solution concepts (see, for that, Kreps/Wilson's (1982) concept of sequential rationality).

[11] We show in section 2.4 that neuroscientific adjustment of the set of constraints and opportunities is somewhat different.

of emotional and automatic brain responses as a part of human decision-making can be used for refutations and improvements of economic models. Adding components of social preferences proofed by neuroimaging to conventional economic models is one example. As a background for such a systematic change in economic modeling we have to explain the role of the brain as a human institution of behavior. Here, we refer to Williamson's (1990) approach of an institutional black box applied in Camerer/Loewenstein/Prelec (2004). From institutional economics we know that we could adjust the objective function or the constraints in the optimization problem of the 'institutional' decision-maker. Section 1 of this article showed that neuroeconomists focused their research on preferences and, thereby, the objective function. One may speculate, if the usual economic approach of changing the set of opportunities may be an alternative to the direct adjustment of preference functions. We will not discuss in this article the very fundamental query of the influence of neuroeconomic results on the appropriateness of mathematical economics as mathematically well behaved objective functions and constraints are required to solve sophisticated decision problems.

2.1 On the Neuroeconomic Proof of Economic Models

Economists have to present useful explanation and prediction of human behavior. Based on the 'as if' assumption as well as on the scientific requirements of logical consistency and predictive power, every conventional economic model of human behavior has to show the uniqueness of the theoretical result. Many neuroeconomists state that with the application of their standard tools economists produce many models and many predictions, "each claiming to be the unique predictor" (Rustichini 2005b, 203). For the explanation of present and future behavior, this is not a satisfying state of the art.

Neuroscientific research proves testable propositions of economic theory. Neuroeconomists believe that the study of neural mechanisms behind individual decisions enable us to find the right and unique decision mechanism for the accurate prediction of behavior and economic outcomes. They assume a given set of constraints and search for the real preferences. The set of constraints define the environmental and informational conditions of decision-making. The neuroscientific research on preferences wants to develop a general model of human preferences confronted with different situations, games and constraints. The role of beliefs as "factual statements concerning states of affairs and causal relations" (Gintis 2004a, 57) in the triad of behavioral parameters, namely preferences, beliefs and opportunities, seems to be unclear and is at least not really discussed in neuroeconomics.[12] But neuroeconomists lay stress on the fact that neurobiological underpinning of brain processes supports effectively the choice of the accurate behavioral economic model.

Here, we want to draw some doubts about the (current) ability of neuroeco-

[12] On the role of beliefs in behavioral economics, cf. Glaeser 2004 and Gintis 2006, and for a first interpretation of beliefs from a neuroeconomic point of view, cf. Camerer 2006.

nomics to develop a general tendency towards unique predictions by the choice of the accurate behavioral model. A useful neuroscientific example is the warm glow of playing a human. Rilling et al. (2002; 2004) showed in fMRI studies that some regions of the brain were not activated whenever subjects were instructed to play a typical Prisoner's Dilemma game with a computer rather then with a human being. Furthermore, playing 'contribution' was less common in the game with the computer even under playing against the same strategy choices and with the same monetary pay-off. This could indicate the difference between the warm glow effect in human cooperation and the standard game theoretic effect of non-cooperative behavior. Caused by structural similarities of the games, we assume from now on that the same holds for the public goods game (cf. Fehr/Camerer 2004). One could now localize the emotion of warm glow. This result could help in specifying the utility function and identifying the feeling of warm glow in other neuroscientific studies. Fehr/Fischbacher/Koesfeld (2005) read the results of Rilling et al. as an emotional endowment of at least some individuals with other-regarding preferences.

Let us just develop an illustrative example of this neuroeconomic reasoning. We discuss the warm glow effect in private provision of public goods. The conventional approach assumes a well behaved utility function $U_i = U_i(x_i, G)$ with x_i as individual i's consumption of the private good x and G as total contributions to the public good. Economists apply the game-theoretic concept of the Nash Equilibrium so that the players do not cooperate. The non-cooperative solution is indicated by $U_i = U_i(x_i, G^n)$. With an additional warm glow c of playing a human, the utility function changes to $U_i = U_i(x_i, G^c)$ and the players cooperate. The warm glow $G^n \rightarrow G^c$ seems to be based on the existence of positive other-regarding, i.e. altruistic preferences. The neuroscientific result of playing a human is transformed into a new economic theory of cooperative behavior under altruistic social preferences in a public goods game. Neuroscientific research apparently proves the testable proposition of the existence of social preferences in public goods games.

Cornes/Sandler (1984; 1994) provide a competing proposition. Individuals have some private benefit from the act of contributing to the supply of public goods. Utility depends on their own contribution g_i and not only on the total amount of public goods G. The corresponding utility function with the 'impure' public good[13] is $U_i = U_i(x_i, G, g_i)$. This preference system could also lead to cooperative solution G^c but the outcome of the game comes from non-cooperative behavior. Furthermore, the model could be applied to charitable giving. This is of special interest because charitable giving seems to have an altruistic component in itself. Donors do not receive direct monetary pay-offs from their donations. G is, then, the sum of individual donations. Applying the standard utility function $U_i = U_i(x_i, G)$ individuals act as altruists and show non-cooperative behavior. They have social preferences but do not like to cooperate. With the warm glow of charitable giving (Andreoni 1989), individuals have some private benefit from the act of contributing to the supply of public goods. Utility depends on their own donation g_i and not only on the total

[13] It is a mixed private-public good.

amount: $U_i = U_i(x_i, G, g_i)$. As far as we can see, neuroscientific inspections show a special reward of playing a human. But competing economic models explain the behavior with and without social preferences and cooperative behavior.

We present this example only for purposes of demonstrating the current structural weakness of neuroeconomic plausibility. We do not claim that persons actually cooperate in a public goods game. Many experimental studies in behavioural economics show that they will not. And we do not neglect other important results of behavioral economics in the field of public goods and charitable giving. For example, in the case of charitable giving gift exchange is another reason for existing social preferences (cf. Falk 2007). It can be separated from warm glow effects of cooperation and should be added to the model. But the main general problem with the neuroscientific approach is the imperfect monitoring of the brain responses. The activation of reward-related areas can be a signal for the realization of non-monetary pure private benefits or altruistic dimensions of the reward. In the first case, the feeling of a reward is not correlated to producing economic benefits or costs for other persons. Prestige and honor, for instance, are feelings related to other persons' economic and social positions but give benefits only to the observed decision-maker. Suchlike emotions drive other-regarding behavior in a very broad sense. We feel prestige, if we believe in the importance of our relative social position and having the action known to others, but we do not mind the improvement or reduction of other persons' benefits.[14] The social outcome of the brain response may be identical under the two cases of emotional altruism and prestige seeking, but the 'motivation' behind the response differs. Do we realize the private benefit of prestige or the direct warm glow of cooperation? Obviously, neuroeconomic research overemphasizes the role of direct social preferences like warm glow of cooperation and partially ignores the potential of the 'as if' approach of rational self interest in presenting an alternative explanation of the *same* outcome. Neuroscientific measurement of the quality of emotions seems to be highly imprecise. It is far away from Edgeworth's idea of hedonic measurement.

In our example, it is not directly specified by fMRI or PET tests, if the activation of the area of reward is correlated with a special factor of self-interest or with altruism. We only see that the cooperative behavior is related to some additional reward. The different explanations of economic theory would now imply the neuroscientific feed back of finding brain sub-areas of reward via prestige and reward via the warm glow of altruistic cooperation. The neuroeconomic solution of the problem is, then, definitively not the simple rejection of self-interest. The neuroscientific registration of emotional selfishness upgrades theories of rational self-interest and excludes the 'as if' assumption. Economic 'as if' research based on pure self-interest cannot be constantly rejected by neuroeconomic reasoning, because conventional economic thinking may find alternative explanations and

---

[14] If prestige is correlated with an intended change in the material or non-material pay-offs of other persons, then prestige seeking implies inequality seeking for the rich and equality-seeking for the poor. This contradicts the widely accepted notion of general inequality aversion (see, for inequality aversion, Fehr/Schmidt 1999). The rich one replaces or compensates the negative emotion of guilt by the positive feeling of prestige.

predictions for the actual human behavior and economic outcome observed in neuroeconomic experiments.

This example shows that the economist has to be careful with economic interpretations of neurobiological experiments. Impure altruism or mixed public-private benefits from public goods provision indicate an additional private benefit of the own contribution to the cooperative result without a direct relationship to other persons whereas social preferences leading to fair behavior imply the conclusion that cooperation in itself is a rewarding behavior. In the first case, we stress an egotistic aspect of emotions, in the latter case a social emotion of behavior and interaction. Neuroscientific research tries to prove testable propositions in economic models but economic modeling may challenge the neuroscientific prove. The relationship between actual neural choice and economic theory is a feed back system. Camerer's (2007, C35) conclusion that "preferences are both the output of a neural choice process, and an input which can be used in economic theory to study responses to changes in prices and wealth" ignores the possibility that economic theory itself provides different testable and sometimes competing inputs for the neuroeconomic understanding of certain neural choices. Neuroscientific design of experiments has to be accurately adjusted, if different economic theories explain the neurobiological data. Whenever neuroeconomists fall short of testing all competing economic explanations of preference functions and behavior, then the presented result of a special motivation, in the example above the existence of social preferences driving cooperation, may be a wrong conclusion and indicates a distorted awareness of positive social preferences in the development of economic theories. If we see more accurate propositions and predictions as a result of this kind of research, is not sure. Neuroeconomic reasoning may rule out interesting and probably accurate theories and improvements on them. In our example, the possibility of private benefit components originated by the emotion of prestige will be ignored, if (neuro)economists belief only in positive social preferences, and the prediction of cooperative behavior by an economic theory developed for new social situations or public policy reform may be misleading because the effect of prestige is missed.

Altogether, this shows clearly that Camerer's (2007, C35) second neuroeconomic claim that "if we understand what variables affect preferences, we can shift preferences and shift behavior (without changing prices or constraints)" is very demanding. Concerning this claim, neuroeconomics is still in its infancy and may confuse or mislead by its low exactness our understanding of preference systems. The contribution of neuroeconomic reasoning to the better development of economic logic may be overstated and the contribution of traditional economic thinking to neuroscientific research on economic problems understated. One essential problem driving that statement is that all types of intangibles and intrinsic motives tend to soften the generation of predictable behavior because they complicate the proof of the uniqueness of preferences and outcomes. Neuroscientific research shows that they exist and that we have to integrate them into economic modeling, but how to do this with the required precision is still an unsolved question.

2.2 The Brain as a Black Box and the Macrotechnology of the Brain

Classical microeconomics is divided into two parts. The first one is the theory of the private household with the household as the consumer of goods and services and the supplier of production factors, the other one the theory of the firm explaining the basics of goods supply and factor demand. The basic structure is the definition of a simple objective function for the household (utility maximization) and the firm (profit maximization) under certain constraints. The preferences of the household are defined by a utility function specified for the different dimensions of household decision-making like the consumption structure, work supply or private saving. Endowments, abilities and market prices for goods and services define the opportunities of private households. Optimization of the preference function subject to these constraints gives the maximized household welfare. The firm combines inputs like labor, capital and land via a production function towards a maximization of the firm's profit. Factor prices, capacity constraints and the productivity of the input factors determine production cost and production opportunities and are the main constraints of profit maximization under given output prices.

Williamson (1990) argued that the traditional neoclassical theory treat the firm and its internal organizational and contractual nature as a black box. He refined and generalized his black box critique towards the neoclassical analysis of different organizations like states, clubs and other contractual relations all dealing with the organization and the optimal degree of vertical and horizontal internal integration of production and exchange (cf. also: Williamson 2002). Institutional and organizational economists examined the theory of firm in more detail and studied the governance structures and internal relations between owners, managers and shareholders as well as the economic rationality of firms, markets and states in vertical and horizontal exchange relationships. As a main result for the theory of the firm, the single objective function of profit maximization as a conception for the 'rational self-interest of the firm' was refuted. The interaction of different members, levels, or, as Camerer (2007, C28) consequently puts it, 'components' of the organization governs firm's final decision and outcome.

Camerer/Loewenstein/Prelec (2004, 556) and Camerer (2007, C28) draw an analogy of the black box conception for the organization of the human mind. Compared with the structure of traditional neoclassical microeconomics, neuroeconomics examines the black box of the private household/consumer whose objective function is directly based on individual preferences. Different organizational components of the mind, i.e. brain areas and brain processes, drive human behavior. And we reported in section 1 of this article that the main result is the refutation of preference functions solely based on the 'rational self-interest of the human being'.

The analytical correspondence of the governance of the firm and the brain is apparently plausible. Nevertheless, we cast doubt on this analogy. Besides the fact that social preferences inducing gift exchange and reciprocity explain a lot of interaction inside the firm (cf. Bellemare/Shearer 2007; Maximiano/Sloof/Sonnemans 2007), the bulk of contributions to the organizational theory of the firm does not show the 'natural' threat to depart from the paradigm of rational self-interest. Organizational structures are investigated with the help of neoclassical analysis. The many models of positive and normative principal-agent theory for the study of the relations between owner and manager are typical examples for this research. The neuroeconomic foundation of the organizational theory of brain and mind need the discussed paradigm shift. Dual-self approaches in economics[15] try to combine different components of the mind, to identify conflicts[16] between them and to prospect for economic mechanisms in human decision making.

Instead of transdisciplinary integration required by the organizational theory of the brain we prefer to adopt an alternative analogy from conflict economics. Hirshleifer (2000) insists that economists should not try to explain or design microtechnological conditions of conflict goods and contests. Technological specification and development of military hardware is the task for technical experts and military or security theorists. Prominent examples of microtechnological factors are the attrition rate of forces in a conflict and the technological conditions of offense and defense (Hirshleifer 2000, 782–787). For Hirshleifer (2000, 774), "[t]he economist's role is not to replace such professionals but instead to address the macrotechnology of conflict, making use of such familiar concepts as increasing versus decreasing returns, economies of scale and scope, and factor substitution".

The microtechnology of the brain is exactly the functioning of different brain areas and the brain processes, and this is, as Camerer (2007) and others put it, the field of neuroeconomic study. From the macro-micro-technology point of view this is not the key area of economic research. We can transfer the mentioned conditions of a production function into the concept of a preference function. We then ask for the rates of substitution in consumption, the convexity of preference systems, the measurement of benefits, rewards and other-regarding components, and so on. As the conflict success function is an abstract analog of the production function under peaceful fabrication, we need some abstract neuroscientific analog of the traditional preference function for a meaningful application of economic tools to the analysis of the brain and its outcome called human behavior. Defining and specifying an economically well behaved human preference function by the means of neuroimaging is different from the black box conception. The first research agenda relies on the rationality assumption whereas the second one allows for irrational behavior and microtechnological 'extra-economic' determinants of behavior. It is not clear how the black box conception may contribute to the feed back from economics to neuroscience, i.e. from the macrotechnological level to the microtechnological level, because it is at least partly based on non-economic components.

---

[15] A short overview is presented in Camerer 2007, C28; cf. also: Sanfey et al 2006. For older interdisciplinary work in the field of the multiple self, see Elster 1986, and for an early approach to the dual- or multiple-self order, see Hayek 1952; 1982.

[16] E.g., the will to be rational and the weakness of will.

2.3 Emotions and Beliefs as Neuroeconomic Constraints

The black box view nevertheless has its merits because its examination implies the perception of a big difference between the solution of the black box problem in organizational economics and neuroeconomics. The repetition of Camerer's (2007, C35) second neuroeconomic claim that "if we understand what variables affect preferences, we can shift preferences and shift behavior (without changing prices or constraints)" shows clearly that the idea of neuroeconomic analysis is restricted on the direct study of preferences. The variation of mind 'constraints' in the sense of restrictions on controlled decision-making is refuted.

The theoretical treatment of asymmetric information in the principal-agent problem between the owner of the firm and the manager illustrates that crucial difference. The owner can not control perfectly whether the manager acts in the interest of the owner. The manager can use his informational advantage to realize an informational rent at the expense of the owner. We can write down the owner's preference for best possible control directly in the preference function. We measure the reward of control, and then we shift preferences and probably the owner's behavior without changing constraints.

It is well known that conventional economists map the principal's problem in a different way. They design the owner's objective function in terms of unrestricted (expected) utility and place the interest of control into the constraints. Incentive compatibility constraints (ICC) and participation constraints ensure the best possible behavior of the manager from the owner's point of view.[17] To solve the problem, we substitute the constraints into the objective function or apply another solution method.

In neuroeconomics as well as bioeconomics, emotions driving benevolence or malevolence change the objective function directly. From the conceptual point of view we can transform this direct effect into constraints. Emotion compatibility constraints (ECC) can ensure that the solution of the decision problem includes the emotional or automatic components of decision-related brain regions. We may formulate conditions for emotional negative or positive rewards or other emotional influence on the objective function. All these constraints have to be based on neuroscientific or other experimental data. This method probably can not simulate neural circuits and other complex neural processes, but it is an abstract analog of the economic theory of constrained optimal decision-making. Neurological opportunities of the brain enter the set of informational and classical feasibility constraints. It is an optimistic and speculative analogy that the conceptual tool of constrained optimization can be effectively applied to the incorporation of neurobiological features of behavior into economic reasoning. But we can improve economic logic by the application of stylized neuroscientific constraints on certain variables and parameters of the objective function. Camerer (2007, C39-C40) implicitly complied with coverage of neural constraints in economic theory. But neuroeconomics as well as conventional economics still fail to

[17] Typically, incentive compatibility leaves some room for informational rents. But the welfare loss of the owner caused by the informational disadvantage is minimized.

create constraints in economic models that cover emotionally based activities, automatic responses and imperfections of the brain.

Specifying neural constraints in problems of economic optimization does not, in principle, exclude an additional neuroeconomic specification of objective functions. Informational problems, for example, imply a specification of objective functions in terms of expected utility. And neuroeconomic problems may require terms of other-regarding preferences. Section 1 of this article showed that ERM structure such an extension. Pesendorfer (2006, 720) is right when he emphasizes that conventional economic modeling "[can not] deal convincingly with the hypothesis that people are wrong about their objective function" but partly misses the point by ignoring that aspects of neural irrationality be left over for neuroscientific constraints. We can, for example, apply the neuroeconomic constraint of nonlinear probability weighting to an objective function the principle of which be still expected utility so that neural processing of probabilities is not completely ignored.

A notional example based on Hirshleifer (1998, 458-459) and referring to beliefs may illustrate the point. In Figure 1 each point represents an income distribution $\{Y_i; Y_j\}$ between two self-interested persons i and j. Preferences are neutral (N) with respect to the position of the opponent so that every individual wants to attain the highest possible income position. Indifference curves (I) for i are vertical lines, for j horizontal lines. The curve EG constraints the feasible set of income distributions. Individuals may now play conflict (defect) or peaceful production and exchange (cooperate).

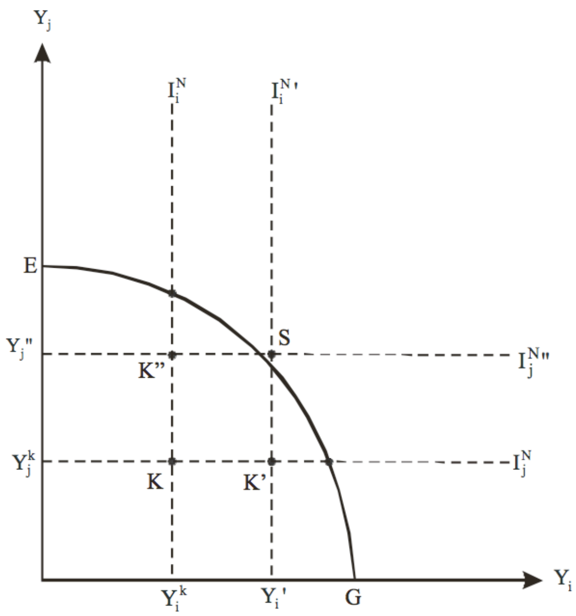

Figure 1: Conflict, exchange and relative optimistic beliefs

The players know that any investment in conflict is costly. Conflict destroys income generating opportunities. The conflict point K shows a typical income distribution under conflict. We would, therefore, expect that the players cooperate and find a solution in the settlement region the area northeastern of K and restricted by EG even if players feel uncertain about the actual conflict outcome. They simply agree on the perceived conflict distribution K.

Now assume that neuroscientific laboratory experiments show that the rate of cooperation is significantly lower than the model predicts, and that playing conflict stimulates a reward-related brain area giving an additional non-material reward to the conflict solution. Each player seems to have an optimistic error of belief concerning the material outcome of conflict.[18] By assumption cost-benefit analysis transfers this non-material reward into private monetary equivalents (Y_i' $- Y_i^k$) and (Y_j" $- Y_j^k$). i perceives K' and j believes in K". The results are the following: First, there is no uniquely expected conflict distribution. Second, each side is relatively optimistic. Third, in the graphical example each side perceives peaceful exchange as a pie in the sky since the different beliefs move the possible settlement region outward from the origin to such an extend that the perceived peaceful cooperation is outside the feasible set of income distributions (see corner S of the Pareto field).

If neuroimaging repeatedly show a significant rate of conflict and an additional reward of conflict choice indicating relative individual optimism and/or overconfidence, then economists should not assume agreed conflict allocations and should not expect cooperation as the 'natural' outcome of the situation. They have to add a neuroeconomic belief constraint (NBC) to the economic problem: "Maximize (expected) utility (objective function) subject to (1) relatively optimistic beliefs/individual overconfidence related to conflict distributions (NBCC) and (2) the feasibility constraint on income distributions. NBC restricts the model to situations of choice in which K" lies to the northwest of K'!'

Whenever neuroeconomic experiments could show NBCC, the additional conflict constraint could have tremendous consequences for economic theory. A unique status quo allocation could not be simply assumed in models in which the status quo is based on conflictuous interactions. An example for such a setting would be the state of nature theory of anarchy, dictatorship and constitutional conflict resolution (cf. Bush/Mayer 1974; Buchanan 1975; Hafer 2006; Acemoglu/Robinson 2006). The NBCC constraint could probably influence any modeling of settlement in the shadow of conflict (see, for an overview, Garfinkel/Skaperdas 2007).

As long as mainstream economic modeling is reasonably dominated by the conception of (constrained) rational behavior (see, for that, Gul/Pesendorfer 2005; Pesendorfer 2006), the addition of neuroscientific constraints to rational economic optimization is an important step towards the integration of neural and conventional aspects of economic choice. And, confronted with Gul/Pesendorfer's (2005, 40) aversion against the excessive neuroscientific bur-

---

[18] On the positive value of systematic errors, cf. also: Waldmann 1994.

den on economic theory, we believe, that this research agenda place a reasonable and characteristic burden.

But this is not the whole story. Referring to section 2.1 of this article, a conventional economist might claim that he found another explanation for the behavior reported in the protocol of the neuroeconomic experiment. The high rate of conflict could be also explained, if we substitute self-interest for a sufficient degree of mutual malevolence. This could imply hostile behavior with a clear preference for the conflict outcome. And this emotion could stimulate the same reward-related brain area giving an additional non-material reward to the conflict solution. Furthermore, an intrinsic motivation for playing conflict could exist leading to intrinsic rewards. The players would, then, simply enjoy playing conflict. But compared with the conventional model, this rather means a change in the preference system instead of a change in the structure of beliefs. Neuroscientific inspection would now be demanded for specifying activities of reward-related brain regions connected to malevolent preferences, intrinsic motivation and/or relatively optimistic beliefs.

Furthermore, the errors of belief which result may be systematic. Additional non-material rewards can be strictly outcome-dependent. In our example, the additional reward of playing conflict is restricted to the case of advantageous conflict distributions. We do not feel relative optimism, if we may be led to a disadvantageous income distribution. In this consequentialist notion human behavior is outcome-dependent. Also, malevolent behavior could be shown only if it 'relatively pays'. On the other hand, malevolent behavior might be restricted to disadvantageous income distributions. A preference system showing inequality aversion (Fehr/Schmidt 1999) is the standard example of outcome-dependent preferences in behavioral economics. Only in situations of disadvantageous distributions individuals feel envy, whereas in situations of advantageous distributions individuals feel guilt. The observed brain activity would contradict the application of this conception of social preferences to our problem of choice.

A related procedural notion is action-dependence (Hirshleifer 1987, 317–321). In our example, the additional reward depends, then, on the fact that the contender plays conflict. Whenever he chooses peaceful exchange, we cannot identify an error of belief with additional non-material reward for cooperative actions. An income distribution that could be tolerabale as a result of peaceful exchange, might lead to an additional brain response if seen to be the result of conflictuous action on the part of the other agent. Relative optimistic beliefs would be the result. The behavioral theory of social preferences also contains a relevant concept for action-dependence. A player shows kindness, if he expects that the opponent is also a friendly person, and he switches to meanness, if he beliefs in the choice of unfriendly actions on the part of the contender (Rabin 1993). But, again, this theory leads to additional non-material pay-offs only in the case of mutual kindness and explains a drift towards contingent cooperation. It does not coincide with the reported brain response.

The discussion of the notational example shows that fruitful coexistence of ERM and neuroeconomics requires definitively permanent feedback. Neuroscientific experiments improve and specify ERMs. In the example, the recorded reward would contradict the ERM standard models of behavioral economics. We could not recommend them for an economic theory explaining the conflict-exchange situation above. But ERM with an error of belief, another kind of malevolence or intrinsic motivation for playing conflict all of which to be defined in the set of constraints or directly in the objective function could lead to an alternative explanation the verification of which would need neuroscientific support. This helps clearly in choosing the relevant economic model by integrating (or ignoring) meaningful beliefs and/or emotions in the model's set of constraints. ERMs put the threat on neuroeconomics to specify neuroscientific laboratory experiments. One example mentioned above is the specification of a belief or an emotion as being outcome-dependent or action-dependent. The 'reply' of neuroeconomic analysis specifies the type of the constraint in ERM.

A continuation of Camerer's (2007) analogy with the organizational economics of principal-agent relationships may be helpful here. In a world of complete contracting, the contract designer specifies fully contingent contracts by fixing type-dependent incentive compatibility constraints and participation constraints. And with asymmetric information, the contractual equilibrium may be a partly separating equilibrium in which some of the different types of agents choose different contracts or a partly pooling equilibrium in which some of the different types choose the same contract. In the same analytical way, human behavior may be contingent on outcome-dependent or action-dependent constraints. If all constraints are binding we see neuroeconomically separated behavior, and if there are some non-binding neuroeconomic constraints, we detect pooled behavior.[19] In the latter case, not all of the neuroscientific conditions which are detected in the lab are of behavioral importance.

Because feed back will always occur, we do not believe in the "generalizability" (Levitt/List 2007, 153) of certain neuroeconomic results for economic theory. And Levitt/List (2007) cast by a somewhat different but related line of argumentation into doubt the extrapolation of neuroeconomic laboratory data to the real world. Indeed, generalizability of neuroeconomic data is a problem for the development of economic models referring to other situations than the one tested in the labs as well as for the application to other real world situations, populations or cultures outside the laboratory. As a consequence, Levitt/List (2007, 154) claim that "interpreting laboratory findings through the lens of theory helps us to understand the observed pattern of results and facilitates extrapolation of lab results to other [theoretical and practical] environments".[20]

3. Towards the Unification of Behavioral Sciences

Section 2 reported, in principle, the fruitful interdependence of neuroeconomics and conventional economic theory and depicted critically some ways of structured integration. As we outlined in the introduction, Smith and Edgeworth proposed a unified theory of human behavior. Whereas Edgeworth restricted

[19] The same analogy may work for incomplete contracting.
[20] Phrase in squared brackets added by the author.

his research on the pure economic man, Smith tried to develop a general theory. Hayek, the third 'classic economist' in the field, was more skeptical about a unified approach based on studies of the brain. Hayek (1952) identified a strict dualism between physics and psychology, and we can extrapolate this dualism without any losses of generality and quality to the dualism between economics and psychology, at least whenever we try to inspect the mind with mathematical economics. Can neuroeconomic analysis solve this dualism, or shall we refer to different analytical systems supporting each other in the best way we can do? We find three different approaches to solve that question:

- Neuroeconomics as a unified theory of human behaviour
- ERM as a unifying approach
- The multiple-system approach of neuroscience

Rustichini (2005b) prefers the first, Gintis (2004b, 2006) believes in the second, and Sanfey et al. (2006) finally propose the third method.[21] Rustichini starts with Smith's (1759) concept of individual sympathy with the emotions of observed individuals. From the viewpoint of ERM he tries to argue that social preferences and prosocial behavior are fundamental to the social sciences and that the neuroeconomic way is the only one to integrate them into a general theory of behavior. The emotion of sympathy leads to a simulation reproducing what a person would feel in the situation of others. This is, from the viewpoint of Rustichini, the essential element of a unified social science. Conventional economics, solely based on self-interested actors, ignores this basic element of human life and behavior in a social environment. Other sciences can not deal convincingly with the selfish motive. Rustichini now argues that neuroeconomics specifies sympathy as an innate attribute of human choice starting from the observation of the others' acts or affective reactions. The most powerful point for the unifying appeal of neuroeconomics is the replacement of psychological introspection by the tools of neuroscientific research. This seems to remove the intellectual limits Hayek (1952) suggested. In neuroscience, it is not our own mind concluding for mechanisms in the mind of human beings but it is a machine documenting brain activities. Neuroeconomic analysis, Rustchini conjectures, provides the unified model for the research agenda of Smith.

It is a first striking argument that mirror neurons and neural mirror systems producing the emotion of sympathy work in the same way when the subject performs emotionally and when the same subject observes others showing the identical emotion. There is a certain consilience between being active "when the action is performed and also when it is observed" (Rustichini 2005b, 208). This leads to neural social understanding by simulation inside the pure observer which allows the observer to internally reproduce and understand the internal state of

[21] Despite the fact that Sanfey et al. 2006 also sympathize with the second approach, they see the usefulness of the unification in economic models only in a reference/benchmark point for advanced neuroscientific research. Examples they mention are the neural bases of reward, value and probability estimation as well as the interfaces between them.

other persons. The actor has not to execute personally the behavior. Rustichini asserts neuroscientific research on the human emotions of disgust and pain the results of which support the proposition, "that there is a substantial overlap between the areas that are activated when we experience an emotion and when we observe someone experiencing that same emotion" (Rustichini 205b, 209).

Ahead of concluding in comparison with the other two approaches, we may place four suggestions on the unifying view of neuroeconomics. At first glance, the identification of neural sympathy seems to be independent of human introspection. Hayek's recognition of the limits of generating meaningful introspective insights which ends in his proposition of practical dualisms is apparently refuted. But, secondly, the interpretation of the recorded brain areas and brain processes is in itself an introspective process of the researcher. Hayek's philosophical statement is not completely disproved. Thirdly, the basic inference assumes implicitly that the mirror systems of human beings are substantially similar. As far as we can see, we need a lot of additional neuroscientific data for a verification of this presumption. And forthly, the neuroeconomic unification is not successful due to lack of information. Whenever we call an approach an economic one we have to deal with preferences, beliefs *and* opportunities. And the unifying momentum of neuroeconomics is focussed on novel preference structures without any scientific balancing and theoretical 'harmonizing' with beliefs and neuroeconomic constraints. The discussion in section 2.3 of this article showed clearly that we have to check also the structure of constraints and beliefs for a neuroscientific updating or adoption of economic models. Last but not least, we can not identify a new unifying paradigm of neuroeconomics. This impression leads us to the multiple-system approach of neuroscience. But before moving to that proposal of differentiation we like to discuss the unifying prospects of ERM.

The neuroeconomic approach is based on the strategy how to unify models by neuroscientific inspection and coherent improvement. In contrast to that the second approach to provision of a unified theory claims that a unified view on behavior should start with the construction of a common underlying model of choice. Then, there is no room for an additional strategy how to unify different models. Gintis (2004b) mentions the preconditions for a unified theory: compatibility, consistency, synergy and enrichment. The most important feature of compatibility in behavioral sciences is the inspection of individual behavior, which is sometimes aggregated in the different disciplines by differing rules towards social behavior. Consistency requires that comparable models are applied in different behavioral sciences whenever they study the same topic. Synergy means the ease of considerable updating of a science by new relevant insights of the neighbor disciplines. The precondition of enrichment differs from synergic unification in that it involves appropriate expansion of the common underlying model in the specific manner every discipline requires.

Despite the listing of different scientific features of unity,[22] the ultimate focus of Gintis (2004b) is on the rational actor model, "a flexible tool that applies to all the human behavioral disciplines". Nowadays, it is applied in economics, political sciences, evolutionary biology, sociology and psychology. Economists

[22] Gintis 2004b; 2006 refer to game theory and evolutionary biology.

apply and enrich rational actor models from political sciences, psychologists are interested in proving different parameters of rationality, and sociologists apply the rational actor model to inspect aspects of social power. An intriguing example of a unifying game-theoretic model with rational actors is the public goods game. Gintis (2004b) cite numerous papers in political science, economics and psychology all of them implying cooperation when costly punishment is permitted. Quervain et al. (2004) specify the neural processing of such a behavior. ERMs can be based on neuroeconomic and bioeconomic insights[23] but may be also expanded by contributions of other behavioral sciences. ERM seems to be compatible, consistent, synergic, and open for various enrichments from different disciplines. Gintis (2004b, 2006) shows that transitive preferences and the option to choose between actions or consequences is all what we need. Whereas Gintis (2004b) complies with the focus of behavioral economics and neuroeconomics on the formation and empirical determination of consistent preferences,[24] Gintis (2006) also considers choice to be contingent upon beliefs and constraints.

Sanfey et al (2006, 109) and Gintis (2006) discuss, mitigate and refute a lot of rooted objections against the modeling strategy of ERM. A common point is the evolutionary stability of the behavioral mechanism of approximately optimization. Then the idea and identification of 'perfect' optimal behavior is a useful benchmark for the measurement of deviations. Expanding the space of preferences, beliefs and/or constraints based on careful investigations of actual behavior in laboratory experiments recovers very often the rational actor model and removes apparent choice inconsistency. Additional or competing 'as if' approaches targeted on the same removal may enforce further neuroscientific studies. Furthermore, Gintis (2006) argues forcefully against the general relevance of the conception of individuals as poor planners. Inaccurate or confused reasoning is, as Gintis (2006) and Hirshleifer (1998) state, a performance error based on the prohibitive costs of perfect education and perfect decision-making at the time of actual behavior. The recourse of the agent to evolutionary stable heuristics like imitation or herding ends in the solution of the (evolutionary) ERM.

The main point for our discussion is that the unification procedures of ERM and neuroeconomics are distinct ways. For the ERM unity, neuroeconomic experiments play a role for the empirical determination of transitive preferences but neuroeconomic data do not form the dominant principle in the ERM conception of unity. Neuroscientific research transforms many 'as if' models into models of actual behavior. Other models are refuted by neuroscience not because they base on the rational actor but because they need a neuroeconomic expansion. That neurosciences currently play a prominent role in the tendency towards unification of behavioral sciences is based on the facts that the brain is

---

[23] For Gintis 2004b, the evolutionary foundation of bioeconomics contributes much more to the unification of behavioral sciences than the neuroeconomic approach because, as we mentioned in section 1 of this article, evolutionary causes lead to final causes of human behavior which directly support the interior logic of the rational actor approach.

[24] For the sake of bring forward the conception and the preference consistency of expected utility, the usual exception is the discussion of beliefs concerning probability formation.

the central human decision making mechanism and that other disciplines apply only very rudimentary, if any 'logic' of cerebral functions.

Referring to the many caveats against the unifying aspects of neuroeconomics, we prefer the unifying power of ERM. The broadening of the rational actor approach may be supported (or seemingly refuted) by neuroeconomic results but is structurally independent from neuroscientific research. The adoption of the rational actor in other behavioral sciences than economics and the synergic and enrichment effects between these disciplines started quite earlier than modern neuroeconomic analysis. For example, consider social preferences and optimistic beliefs in conflict outcome. One may integrate these features in the conventional model of rational choice without any recourse to neuroscientific data, and, then, neuroeconomic verification, specification or refutation may help us in enhancing the model. Neuroeconomic experiments shed light on new determinants and parameters of the preferences, beliefs and constraints in rational actor models ignored in previous research. But neuroeconomics seems not to be the principle engine of a unification of behavioral sciences.

The multiple-system approach, as compared with the above proposals, "challenge[s] the core assumption in economics that behavior can be understood in terms of unitary evaluative and decision-making systems" (Sanfey et al. 2006, 111). In contrast to Rustichini's neuroeconomic suggestion, the idea of multiple systems originates from the psychological finding of occasionally conflicting interactions between different systems of decision-making. Sanfey et al. (2006) mention dual processing with differences between automatic (emotional) and controlled (deliberative) processes. The evolutionary elimination of performance errors seems to play a role only in the case of automatic, i.e. heuristic-based processes. And the introspective approach is restricted on the registration and description of controlled processes.

The human coordination of the systems is explained by a two-tiered process. Controlled processes watch and contingently overrule the performance of the automatic system. But Sanfey et al. (2006) state that this system hierarchy is not perfect. Another important result of neuroscientific studies mentioned in former sections of this article is that different decision-relevant systems appear to rely consistently on different brain regions. Recent economic studies on dual processing briefly reviewed in Camerer (2007, C28) and Sanfey et al (2006, 112, 114) suggest that the multiple-system approach can be the dominant unifying force in the behavioral sciences. Whereas Sanfey et al. argue that the identification of dual-system decision-making challenge the rational actor approach, Gintis and the author of this article believe in the performance and flexibility of ERM. All what we would need is a 'rationalization' of dual-processing, the specification of the superiority of different processes in terms of cost and productivity[25] and some arguments for the bioeconomic fitness of multiple systems. If these requirements go beyond the scope of ERM, is a task for future research.

[25] Cf. Sanfey et al. 2006, 111: "Automatic processes are fast and efficient, ... but highly specialized ... and therefore relatively inflexible" whereas "controlled processes are highly flexible ... but relatively slow". This refers partly to the economic trade off between rules and discretion or the central terms of the transaction cost approach to organizational economics.

4. Conclusions and Perspectives

Neuroeconomics rediscovers the research agenda of the classical contributions to the economic theory of human behavior. It provides a substantial improvement of the classical mind-related economic approaches by adding and specifying important cerebrally-operated components of human decision-making and by arguing on the unity of behavioral sciences. The reasonable and structured introduction of emotions and other affective components to economic analysis enhance the predictive power of economic studies and draw the economist's attention to the brain processes of human decision-making. Neuroscientific studies recognize structural components of human behavior many of which are ignored in standard economics.

Section 1 showed that reward-related brain areas connect brains responses to the measurement of utility and human behavior. Standard economic measurement of benefits and cost nevertheless illustrates that neuroscientific measurement is very imprecise. Furthermore, we suggest a deeper neuroeconomic examination of the early economic concepts of intangibles and intrinsic motivation. Conventional cost-benefit analysis could not deal convincingly with these types of benefit and cost. It seems that neuroeoconomics can provide an essential contribution to the modern theory and practice of cost-benefit analysis. Up to now, neuroeconomic experiments are based on very simple models like the classical prisoner's dilemma game, the ultimatum game or the dictator game. Despite the explanatory power of new neuroscientific results concerning actual behavior of agents confronted within the strategic environment of the particular game, neuroeconomics will attract a great deal of attention, if it can deal convincingly with sophisticated game-theoretic models. This may open the door for neuroeconomic domains other then the inspection of prosocial behavior. The notational example in section 2.3 of this article could be the starting point for the neuroeconomics of conflict with the main emphasis on antisocial behavior.

Nevertheless, the neuroscientific inspection of prosocial aspects like sympathy and other positive social preferences pushed the ERM approach essentially forward. The expansion towards social preferences indicated by behavioral economics and neuroeconomics is an elementary improvement, since this change in the structure of preferences, beliefs and probably constraints enables the rational actor approach to cope with neuroeconomic insights. Even if ERM can not include all kind of apparent human irrationality, it may integrate a lot of neuroscientific insights and otherwise provide the theoretical background for further neuroeconomic studies. Together with bioeconomic studies, neuroeconomics is the fundamental approach to the discovery of the final causes of human behavior.

But as we showed in section 2.1 the relationship between neuroeconomic data and economic modeling is not a one-way street. The choice of the accurate model with the correct objective function and meaningful beliefs and constraints depends on the feedback between neuroscientific data and expanded rational actor modeling. And the analogy of the macrotechnology and the microtechnology of the brain suggested a special division of work between neuroeconomics and economic theory, which differs from the black box analogy. The presumably most controversial result of this article is the request for the neuroscientific improvement of economic models via ECC and NBC. We believe that this auxiliary construction improves the tractability of the neuroscientific adjustment and the adjusted economic model tremendously. In comparison with the direct modification of preferences and objective functions, we should add weight to the creation of outcome- or action-dependent neuroeconomic constraints for ERM. And exactly these properties of ERM lead us to the conclusion that ERM is the first candidate for the unification of behavioral sciences.

Bibliography

Acemoglu, D./J. A.Robinson (2006), *Economic Origins of Dictatorship and Democracy*, Cambridge et al.

Andreoni, J. (1989), Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving, in: *Economic Journal 100*, 464–477

Becker, G. (1976), Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology, in: *Journal of Economic Literature 14*, 817–828

Bellemare, C./B. Shearer (2007), *Gift Exchange within a Firm: Evidence from a Field Experiment*, Institute for the Study of Labor, IZA Discussion Paper No. 2696, Bonn

Boadway, R. W./D. E. Wildasin (1984), *Public Sector Economics*, 2nd Ed., Boston-Toronto

Bowles, S. (2007), *Social Preferences and Public Economics: Are Good Laws a Substitute for Good Citizens?* Università degli Studi di Siena, Dipartimento Di Economia Politica, Discussion Paper N. 496, Siena.

Buchanan, J. M. (1975), The Limits of Liberty, Chicago-London

Bush, W. C./L. S.Mayer (1974), Some Implications of Anarchy for the Distribution of Property, in: *Journal of Economic Theory 8*, 401–412

Camerer, C. F. (2006), Wanting, Liking, and Learning: Neuroscience and Paternalism, in: *University of Chicago Law Review 73*, 87–110

— (2007), Neuroeconomics: Using Neuroscience to Make Economic Predictions, in: *Economic Journal 117*, C26–C42

—/G. Loewenstein/D. Prelec (2005), Neuroeconomics: How Neuroscience Can Inform Economics, in: *Journal of Economic Literature 42*, 9–64

Colander, D. (2007), Edgeworth's Hedonimeter and the Quest to Measure Utility, in: *Journal of Economic Perspectives 21*, 215–225

Cornes, R./T. Sandler (1984), Easy Riders, Joint Production, and Public Goods, in: *Economic Journal 94*, 580–596

—/— (1994), The Comparative Static Properties of the Impure Public Good Model, in: *Journal of Public Economics 54*, 403–421

Darby, M. R./E. Karni (1973), Free Competition and the Optimal Amount of Fraud, in: *Journal of Law and Economics 16*, 67–86.

Edgeworth, F. (1981), *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*, London

Elster, J. (ed.) (1986), *The Multiple Self*, Cambridge et al.

Falk, A. (2007), Gift Exchange in the Field, in: *Econometrica 75*, 1501–1511

Fehr, E./C. F. Camerer (2004), Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists, in: J. Henrich/ R. Boyd/S. Bowles/C. Camerer/E. Fehr/H. Gintis (eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*, Oxford,

55–96
—/K.Schmidt (1999), A Theory of Fairness, Competition, and Cooperation, in: *Quarterly Journal of Economics 114*, 817–868
Garfinkel, M. R./S. Skaperdas (2007), Economics of Conflict: An Overview, in: T. Sandler/K. Hartley (eds.), *Handbook of Defense Economics, Volume 2: Defense in a Globalized World*, Amsterdam et al., 649–709
Gintis, H. (2004a), The Genetic Side of Gene-Culture Coevolution: Internalization of Norms and Prosocial Emotions, in: *Journal of Economic Behavior and Organization 53*, 57–67
— (2004b), Towards a Unity of the Human Behavioral Sciences, in: Politics, *Philosophy and Economics 3*, 37–57
— (2006), *A Framework for the Unification of the Behavioral Sciences*, University of Massachusetts, Mimeo, Northampton/MA (to be published in: *Behavioral and Brain Sciences*)
Glaeser, E. L. (2004), Psychology and the Market, in: *American Economic Review 94*, 408-413
Gul, F./W. Pesendorfer (2005), *The Case for Mindless Economics*, Princeton University, Mimeo, Princeton 2005
Hafer, C. (2006), On the Origins of Property Rights: Conflict and Production in the State of Nature, in: *Review of Economic Studies 73*,119–143
Harbaugh, W. T. (1998), What do Donations Buy? A Model of Philanthropy Based on Prestige and Warm Glow, in: *Journal of Public Economics 67*, 269–284
Hayek, F. A. (1952), *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*, London-Henley
— (1982), The Sensory Order after 25 Years, in: W. B. Weimer/D. S.Palermo (eds.), *Cognition and the Symbolic Processes*, Vol. 2 (based on the second Pennsylvania State University of Conference on Cognition and the Symbolic Processes held May 16–19, 1977), New York, 287–293
Hirshleifer, J. (1987), On the Emotions as Guarantors of Threats and Promises, in: Duprè, J. (ed.), *The Latest and the Best: Essays on Evolution and Optimality*, Cambridge/Mass., 307–326
— (1998), The Bioeconomic Causes of War, in: *Managerial and Decision Economics 19*, 457–466
— (2000), The Macrotechnology of Conflict, in: *Journal of Conflict Resolution 44*, 773–792
Kenning, P./H. Plassmann (2005), NeuroEconomics: An Overview from an Economic Perspective, in: *Brain Research Bulletin 67*, 343–354
Kenning, P./H. Plassmann/D. Ahlert (2007), Application of Neuroimaging Techniques to Marketing Research, in: *Qualitative Market Research 10*, 135–152
Kreps, D. M./R. Wilson (1982), Sequential Equilibria, in: *Econometrica 50*, 863–894
Levitt, S. D./J. A. List (2007), What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? In: *Journal of Economic Perspectives 21*, 153–174
Maximiano, S./R. Sloof/J. Sonnemans (2007), Gift Exchange in a Multi-Worker Firm, in: *Economic Journal 117*, 1025–1050
Montague, P. R./G. S.Berns (2002), Neural Economics and the Biological Substrates of Valuation, in: *Neuron 36*, 265–284.
—/B. King-Casas/J. D.Cohen (2006), Imaging Valuation Models in Human Choice, in: *Annual Review of Neuroscience 29*, 417–448
Nelson, P. (1970), Advertising as Information, in: *Journal of Political Economy 81*,

729–754

Pesendorfer, W. (2006), Behavioral Economics Comes of Age: A Review Essay on *Advances in Behavioral Economics*, in: J*ournal of Economic Literature 44*, 712–721

Rabin, M. (1993), Incorporating Fairness into Game Theory and Economics, in: *American Economic Review 83*, 1281–1302

Rilling, J. K./D. A. Gutman/T. R. Zeh/G. Pagnoni/G. S. Berns/C. D. Kilts (2002), A Neural Basis for Social Cooperation, in: *Neuron 35*, 395–405

—/A. G. Sanfey/J. A. Aronson/L. E. Nystrom/J. D. Cohen (2004), Opposing Bold Responses to Reciprocated and Unreciprocated Altruism, in Putative Reward Pathways, in: *Neuroreport 15*, 2539–2543

Rustichini, A. (2005a), Emotion and Reason in Making Decisions, in: *Science 310*, 1624–1625

— (2005b), Neuroeconomics: Present and Future, in: *Games and Economic Behavior 52*, 201–212

Sanfey, A. G./G. Loewenstein/S. M. McClure/J. D.Cohen (2006), Neuroeconomics: Cross-Currents in Research on Decision-Making, in: *Trends in Cognitive Sciences 10*, 108–116

Sen, A.K. (1977), Rational Fools: A. Critique of the Behavioral Foundations of Economic Theory, in: *Philosophy & Public Affairs 6*, 317–344

Shiv, B./G. Loewenstein/A. Bechara (2005), The Dark Side of Emotions?, in: *Cognitive Brain Research 23*, 85–92

Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*, London-Edinburgh

— (1759), *The Theory of Moral Sentiments*, London

Smith, T./A. Tasnádi (2003), *A Theory of Natural Addiction*, Global Fellows Working Paper, International Institute of the University of California, Los Angeles

Waldman, M. (1994), Systematic Errors and the Theory of Natural Selection, in: *American Economic Review 84*, 482–497

Williamson, O. E. (1990), A Comparison of Alternative Approaches to Economic Organization, in: *Journal Institutional and Theoretical Economics 146*, 61–75

— (2002), The Theory of the Firm as Governance Structure: From Choice to Contract, in: *Journal of Economic Perspectives 16*, 171–195

Zak, P. (2004), Neuroeconomics, in: *Philosophical Transactions of the Royal Society of London, Biological Sciences 359*, 1737–1748