Journal of Neuroscience, Psychology, and Economics
2012, Vol. 5, No. 1, 38–61

1937-321X/12/$12.00 DOI: 10.1037/a0026882

Towards an Externalist Neuroeconomics: Dual Selves, Signs, and Choice

Carsten Herrmann-Pillath
Frankfurt School of Finance and Management

Recently, reductionist models of choice have been proposed which directly reduce a modified version of economic utility theory to neuroscience. I propose an alternative conceptual framework that adopts the position of externalism, which I further narrow down to a distributed cognition framework, and eventually to a biosemiotic model of human choice. I relate this to existing modular theories on the brain, which I generalize into a dual selves model. In this model, the internal information asymmetries and deficiencies of the brain/body system drive the reliance of choices on external cognitive mechanisms as manifested in sign systems. I analyze this interaction and apply the combined biosemiotics/dual selves model to an important set of falsifiers of an integrated reductionist model of neuroeconomics, that is, addiction and other consumption disorders.

Keywords: foundations of neuroeconomics, semiotics, dual selves, distributed cognition, consumption disorders

I. Externalism Versus Internalism in Neuroeconomics

One of the fundamental issues in current neuroeconomic research is the question of whether the analysis of the brain should follow the standard economic model or whether the introduction of an alternative framework is needed. Following the standard economic model means referring to the utility function framework and identifying neural correlates. This paradigm has been established canonically by Glimcher (2009, 2011). Alternative approaches would take certain insights from psychology and the brain sciences as a starting point, such as the distinction between conscious and automatic processes, and would thus result in models which differ from the standard economic model (Camerer, 2007; Camerer, Loewenstein, & Prelec, 2005). The drawback of these views is that they still lack a coherent and consistent theoretical framework comparable to the standard economic approach in terms of its simplicity, consistency and empirical tractability. This deficiency also affects the applied fields of neuroeconomics (as recognized by Hubert, 2010; Hubert & Kenning, 2008). On the other hand, however, a paradox lurks behind the reductionist program. Ultimately, applied (neuro)-economics would no longer need to refer to neuroscience at all, because on the level of observable behavior the (possibly modified) standard economic tools would suffice. It is precisely if the reduction program can succeed that neuroeconomics will become irrelevant for economics because all the essential insights of neuroeconomics would be reproducible within economics without any reference to neuronal facts. This would be the case as long as the phenomena to be explained remain on the ontological level of the units of choice, that is, the individuals (for related criticisms of neuroeconomics, see Bernheim, 2009; Gul & Pesendorfer, 2008; Harrison, 2008). Therefore, Glimcher's (2011) method-

I received important advice from the JNPE referees, which helped to improve the paper substantially. Parts of earlier versions were presented to workshops at Durham University, Jena University and the annual Neuropsychoeconomics conference 2011 and the EAEPE 2011 annual conference at Vienna. I am indebted to Don Ross for his detailed comments. The usual disclaimer applies. Finally, special thanks to Judith Mader who polished my English language.

Correspondence concerning this article should be addressed to Carsten Herrmann-Pillath, Academic Director, East-West Centre for Business Studies and Cultural Science, Frankfurt School of Finance and Management, Sonnemannstraße 9-11, 60314 Frankfurt am Main, Germany. E-mail: c.herrmann-pillath@fs.de

ological claims deserve close attention, and are certainly a crucible for testing the relation between neuroscience and economics.

The Pivot: Externalism in the Philosophy of Mind

In this article my central concern is how to design a conceptual framework for neuroconomics which offers an alternative to the reductionist approach, and *because of this* will demonstrate the ultimate relevance of neuroeconomics and neuroscience for economics. In doing this, I mainly adopt a philosophical perspective, especially the philosophy of mind, in order to construct a conceptual framework which might help to organize the different theoretical propositions and hypotheses that come together at the neuroeconomics interface, although originally emerging from different disciplines (compare Foxall, 2008).

One foundational issue in the philosophy of mind is the distinction between externalism and internalism regarding mental phenomena (for surveys, see Lau & Deutsch, 2008; Schantz, 2004; Wilson, 2004). Reductionist neuroeconomics seems ambivalent with regard to this distinction and this reveals the need for a conceptual clarification. The reason for this is that the standard economic model is externalist, whereas its proposed neuroeconomic version is internalist. Externalism in economics is reflected in the revealed preference approach to the utility function: The utility function refers exclusively to observable behavior with no reference whatsoever to internal mental states; the concept of utility is merely a mathematical device to describe observed actions (Ross, 2005). In contrast, reductionist interpretations of the utility concept refer exclusively to internal neuronal states. So the important difference between economics and reductionist neuroeconomics is that the former is externalist, the latter internalist: This means that reductionist claims actually result in a metamorphosis of the (implicit) economic position of externalism into one of internalism. As a consequence, reductionist neuroeconomics entails a fundamental clash between two philosophical positions implicit in the two disciplinary approaches. I argue that this inconsistency in the philosophical foundations is a sufficient reason to discard reductionism and to search for an alternative that needs to be *externalist* if it aims at an integration between neuroscience and economics. Therefore, my article aims to offer a conceptual framework for *externalist neuroeconomics* (a research program inaugurated by Ross, 2005).

Rendering Externalism Empirically Tractable: Biosemiotics

To achieve the aim of establishing the case for externalist neuroeconomics, I rely on other approaches to externalism in the cognitive sciences which focus on the role of external causal processes establishing mental phenomena in terms of interactions between neuronal states and external facts. These approaches are to be found in the growing literature on "distributed cognition" or the "extended mind" (e.g., Clark, 2011; Hutchins, 1995; Hutchins, 2005; Sterelny, 2004; Ross, Spurrett, Kincaid, & Stephens, 2007; for a critical view, see, e.g., Sprevak, 2009). In this literature, the operations of the human mind are seen as being fundamentally dependent on external facts, referring not only to the obvious role of devices which leverage cognitive skills, but more fundamentally to the externalization of all cognitive processes in the sense that cognitive performance essentially involves the external world and includes both physical entities and social interaction. Therefore, a straightforward sketch of an externalist approach that is different from the externalism of standard economics is possible. This would offer a fresh view on neuroeconomics (for a related argument, see Wilcox, 2008). In this view, if one continued to use the term "utility" in explaining human choice, "utility" would not correspond exclusively to a neuronal state in a neuro-reductionist approach, but to a *causal conjunction between a neuronal state and an external fact*.

In order to render the philosophical approach of externalism applicable in the context of neuroeconomics, I propose to introduce the approach of biosemiotics (which is a new way of making use of semiotics in our context; for surveys on other uses in consumer research, see Mick, 1986; Mick, Burroughs, Hetzel, & Brannen, 2004). According to this view, all mental phenomena are seen as interactions between neuronal states and signs, with signs being ex-

ternal physical facts. These interactions have a *function*, which basically boils down to biological categories such as homeostasis or adaptation, which are relative to the environment. This reflects the etiological approach to function, which sees functions as the result of past and current processes of selection in particular environments (for an overview of theories of function, see Krohs & Kroes, 2009). Thus, the conjunction between internal neuronal states and external facts has a function relative to a context.

This reference to functions implies the rejection of purely internalist conceptions of mental states. In the philosophy of mind, this position has been developed in the context of teleosemantics (Macdonald & Papineau 2006a; Millikan, 1989; Neander, 2009). Teleosemantics claims that all mental phenomena, and most of all the category of meaning, can be explained in terms of functions which emerge from a process of selection. The central problem in this approach is the question of malfunctioning, that is, the dysjunction between certain behavior, based on external information inputs, and the required functioning. Basically, this problem can be rendered in biosemiotic terms: If a frog perceives a fly, it perceives certain movements, shapes, or shades of colors, which are signs that work via perceptions and generate certain responses. These responses can fail if the real object is not a fly (Neander, 2006). Interestingly, in economics, this role of malfunctioning has corresponded to the role of anomalies in driving research in behavioral economics, and ultimately neuroeconomics, in recent decades.

The Proof of the Methodological Pudding: Consumption Disorders and Information Processing in the Brain

Indeed, the acid test for neuroeconomic theories of consumer choice is the entire range of behavioral disorders, such as eating disorders or addictions. These clearly involve the perception of malfunctioning on the part of individuals and their social environment. Addictions are seen as dysfunctions because they can impair proper functioning either in the general sense of health (the smoker suffering from lung cancer) or the more specific sense of fulfilling certain social roles (the gambler squandering the family assets; West, 2006). I, therefore, use these phenomena to empirically substantiate my argument here. This also seems methodologically appropriate as the phenomenon of addiction can be seen as one of the possible strong falsifiers of standard economic models of consumer choice, seminally applied by Becker and Murphy (1988) to addiction (see Ross, Sharp, Vuchinich, & Spurrett, 2008; West, 2006).

Biosemiotics builds on the philosophy of Charles Sanders Peirce (for a selection of important contributions, see Peirce, 1992, 1998, and for surveys on Peirce see Atkin, 2009; Burch, 2010). It emphasizes the generation of information in the coevolution of biological mechanisms such as neuronal processes and signs which are part of the physical environment. This results in certain interpretations of environmental inputs which are manifest in the behavioral responses as outputs. Correspondingly, I argue that neuroeconomics has to build on the interaction between signs and neuronal mechanisms. These can be dysfunctional, depending on whether the generated information leads to responses that fail to meet certain selective criteria in the environment. I, therefore, argue that, for example, addictions to smoking emerge from the interaction between neuropharmacological phenomena on the one hand and environmental determinants on the other, with the latter working via the intermediation of signs. This results in responses which are functional relative to certain contexts. So, within certain contexts, smoking can be functional and would not be classified as an addiction in the clinical sense. This is the case in certain historical periods and demonstrates the difficulties of demarcating the conceptual boundaries of the term. The definitions range from the narrow neuropharmacological to extremely broad definitions claiming that almost all goods and all behavior can become the object of addictions (e.g., Ascoli & McCabe, 2006). The facts about these phenomena are well known and so I am, therefore, not presenting any new empirical insights here. My task is to put these facts into a coherent conceptual framework of neuroeconomics which includes malfunctioning as a special yet methodologically crucial case (in the sense of *experimentum crucis*), and serves to systematize neuroeconomics so as to offer an externalist alternative to reductionist neuro-internalism.

In the most general sense, dysfunctional choice can be interpreted as a failure in information processing, ranging from insufficient external information to the manifold ways in which the internal decision making process might fail (Redish, Jensen, & Johnson, 2008). Let us assume that there is a process in the brain that involves the making of choices, both in terms of conscious operations such as deliberation or in terms of habitual responses. Choice relies on information processing. In the standard economic argument there has been a great deal of emphasis on information about the environment. This information can be imperfect, incomplete, or even false. There is, however, another information problem which emerges in the relation between the decision making process and the organism in toto: information processing with reference to internal states of the organism can also be incomplete, imperfect, and false. As we shall see, this argument even applies to the internal organization of the brain. These fundamental information asymmetries imply that the minimalistic theoretical model of the individual cannot be a unitary one, but must be at least a dual one, which matches the principal-agent models in economics (Brocas & Carrillo, 2008). Thus, the complete framework that I am going to present combines the literature on dual selves models with the biosemiotic approach, resulting in an externalist approach to neuroeconomics.

To summarize, my argument boils down to the following hypothesis (inspired by Don Ross, personal communication). So far, reductionist neuroeconomics focuses on stimulus-centered modeling of valorized learning, which, however, appears to marginalize important recent developments in cognitive sciences, psychology and brain research (this gap is acknowledged implicitly by Glimcher, 2011 when he discusses the unresolved problem of the representation of subjective value and probability). Therefore, I confront this model with a model in which stimuli are seen as being cognitively integrated and mediated via signs, with signs being external phenomena. This is the defining feature of the program of *externalist neuroeconomics*. From this vantage point, the article proceeds as follows. In section 2 I develop the dual selves approach, which underlies the argument of the internal incompleteness of information processing in the brain. Section 3 presents the basics of biosemiotics. Section 4 links the dual selves model with the biosemiotic framework and offers the complete externalist model of neuroeconomics. Section 5 applies this approach to a methodological case study, the economic model of consumption and addiction proposed by Laibson. Section 6 extends this analysis into a general biosemiotic framework for addiction theories. Section 7 concludes by highlighting the significance of the externalist view for the relationship between neuroeconomics and economics.

II. Dual Selves: The Incompleteness of Internal Information Processing in the Brain

Analytically, the dual-selves model is an axiomatic alternative to the standard model of the individual in economics (for an early discussion of alternatives, see Elster, 1986). This means that I do not posit the model as an empirical, hence neuroanatomical hypothesis about brain structure but as a conceptual framework to systematize hypotheses about empirical data. This corresponds to the distinction between "molar" and "molecular" levels of analysis in neuroeconomics: The dual selves model is a most parsimonious molar theory of brain functioning and does not make any specific claim about molecular reduction, while maintaining the general approach that all molar mechanisms must be physically embodied in molecural mechanisms (Ross et al., 2008). It is the most simple, and dual, version of a larger set of theories about brain modularity. As such, the most simple dual selves model could be further enriched by more specific models of (sub)systems, such as a long-term planning system and a short-term impulsive system in the theory of hyperbolic discounting. I would treat these as subsystems of a more general and comprehensive decision-making system, while maintaining the clear distinction between these subsystems as molar-level concepts and the underlying neuroanatomy, that is, the molecular level analysis (e.g., Laibson, 2002; McClure, Ericson, Laibson, Loewenstein, & Cohen, 2007; and the corresponding critique in Ross et al., 2008).

Modularization of the Brain and Incompleteness of Internal Information Processing

The dual selves model must be seen against the background of evolutionary approaches to the individual which argue that the brain is not a fully integrated, coherent and consistent decision apparatus but consists of different modules which are adapted to different tasks which have emerged in past selective contexts (Tooby & Cosmides, 2005; Platt & Padoa-Schioppa, 2009). From this it follows that the coordination between modules is a primordial metaprocess in the brain. The central theoretical argument underlying these modular models is that this coordination cannot be achieved by means of a centralized unit of information processing and decision making (the "homunculus" according to Dennett, 1991).

The theoretical foundation of the dual selves model lies in the principled constraints of information processing in complex neuronal systems. These result from the internal specialization of different parts. The most simple pattern of specialization is the minimal duality of a receptor system as well as effector system. These are connected by means of a mapping relation. The dual selves approach can in this way be introduced as an axiomatic proposition, stating that in a neuronal system with the minimal distinction between a receptor system and an effector system connected by means of mappings, that is, internal representations, mutual information asymmetry between the subsystems is a generic feature of the system. That is, we can apply what is fundamentally an economic argument to the impossibility of complete aggregation and centralization of information in the brain (as argued in Brocas & Carrillo, 2008). This relates to first, the purely quantitative limitations to connectivity, second, to the resulting qualitative differences of functional perspectives, and third, to the role of entropic forces in information transmission (Kåhre, 2002).

These arguments can be further strengthened if we consider the fact that the brain is a complex system of parallel and competing neuronal processes which manifests a dynamism that can be described by models from complexity theory and synergetics (West, 2006; Oullier, Kirman & Kelso, 2008). This can be observed, for example, in the basic distinction between conscious processes and all other processes in the brain, where consciousness is a scarce resource for which a multitude of other processes compete (Dennett, 1991; Calvin, 1998). It follows from this that there are fundamental limits to the formation of a perfect and consistent integrated decision system. There is a direct correspondence with established axiomatic approaches on collective decision-making in economics: Once we suppose the brain to be a system of partly autonomous modules, the Arrow-Condorcet theorem applies, that is, the individual cannot have a consistent and unique preference function (Steedman & Krause, 1986; Ross, 2005).

One of the first systematic, albeit neglected, explorations of the dual selves approach was presented by the sociologist James Coleman (1990), who distinguishes between "Object Self" and "Acting Self," a terminology that I will adopt in what follows. As such, this is a purely theoretical distinction in terms of two different aspects or functions in individual choice, with no presumptions about brain locations. The distinction is rather about two fundamental functional modes in which the brain can operate. One function relates to the evaluation of outcomes of actions (Coleman calls this the "receptor function"), the other function relates to the anticipation of outcomes and the resulting process of choice (Coleman's "actuator function"). If we follow an evolutionary argument, Coleman's proposal can be directly related to the duality of a receptor and an effector system. In this case, the evaluative dimension inheres the past selection of certain combinations between receptor and effector systems through time, without the necessary requirement of a specialized evaluative system intermediating between the two.

A Standard Model: "Wanting" and "Liking"

Dual selves models have a long tradition in both behavioral economics and the related neurosciences, with varying foci. In one strand, the distinction between "experience utility" and "decision utility" has been introduced (Kahnemann & Krueger, 2006; Kahnemann, Wakker, & Sarin, 1997), which would relate to the Object Self and the Acting Self respectively: Individuals manifest systematic gaps between what

they experience as an outcome during certain actions and what they remember and anticipate, even under very simple experimental situations such as feeling the temperature of water. This implies that choices are taken on the basis of information that does not unequivocally reflect the outcomes. This model of dual utility measures differs fundamentally from the standard economic model. The standard economic model focuses only on decision utility, especially in the context of the theory of revealed preference, taking observed behavior as directly representing utilities. This distinction is of relevance in many contexts, such as, for example, in the analysis of the psychological phenomena of self-deception (Mijovic-Prelec & Prelec, 2010).

In the neuroeconomics context, a related fundamental distinction has been made in terms of the "Wanting" versus the "Liking systems." Sometimes a "Learning System" is also introduced, as in Camerer (2006) or Berridge (2009), but, in my reading of the neuroeconomic literature, learning is either a process that takes place in the Wanting system, which builds on dopaminergic circuits enabling operant conditioning, or is an emergent feature of the interactions between the different systems. Separating a Learning system from other systems of the brain seems redundant, as learning is a major mode of operation of the brain as a supersystem. The distinction between "Wanting" and "Liking" results from the identification of specific neurophysiological circuits which enable reinforcement learning, and which mainly involve the neurotransmitter dopamine (for a survey, see Schultz, 2009). The dopaminergic circuits build on the perceived discrepancy between goal and goal attainment, that is, the deviations between expectations and outcomes (i.e., coding reward prediction error). The central point is that perceived well-being does not only depend on the actual outcome, but also on the general changes and levels of dopamine in the brain. This introduces a systematic differentiation between the process of choice and the evaluation of outcomes, and this fits into the Object Self and Acting Self dualism. The dopaminergic circuits are associated with the latter, and all other systems that indicate proper organismic functioning, with the former. This functional differentiation is also reflected in brain architecture, which provides the physical structure in which the dynamic mappings proceed continuously (e.g., Knutson & Wimmer, 2007; Knutson, Delgado, & Phillips, 2009).

However, neuroeconomic knowledge is still in flux, so that the "Wanting" versus "Liking" duality clearly counts as a molar analysis, though building on a number of molecular-level observations. The notion of modularity is complex and, for example, might also include a functional differentiation between systems processing positive (rewards) and negative (losses) outcomes and the signals related to them. This is important for providing a neuroscientific foundation for prospect theory (e.g., Knutson & Greer, 2008; Stanton & Welpe, 2010; Trepel et al., 2005; for an alternative view, see Tom et al., 2007). Another neuroeconomic version of a dual selves approach is the explanation of hyperbolic time preferences in terms of the interaction between two brain modules with different discounting mechanisms (Jamison & Wegner, 2010; McClure et al., 2007). On the contrary, Glimcher (2003, 2009, 2011) sees the brain as a unified decision system as far as the mechanism of choice is concerned (which relates to the common currency hypothesis, see Landreth & Bickle, 2008). Correspondingly, many economists emphasize this position in the neuroeconomics literature and explicitly reject multiple selves models (e.g., Bernheim & Rangel, 2004; in contrast to economic dual selves models such as proposed by Bénabou & Tirole, 2002 or Fudenberg & Levine, 2006). However, both strands in the literature seem to converge on the functional duality of decision making and ultimate evaluation. For example, Glimcher (2009) clearly states that his unified model is a model of choice, but not a model of experienced well-being (which would maintain, among others, the emerging divide between the economic theory of choice and the economic analysis of welfare, see Camerer, 2006, and the cautious remarks in Glimcher, 2011).

Open Ends in the Argument of the Reductionist Alternative

In this context, we must consider that even the most systematic exposition of the integrated model of neuroeconomics presented by Glimcher (2011) has several open ends which indicate the possible need to argue in terms of a dualistic model. This is because choice is primarily seen as being driven by expected subjec-

tive value as well as by the corresponding deviations between realized value and expectations. Following the reward prediction error theory, the values underlying choices are grounded in the difference between what is called "experienced value" *ExperSV* and the discounted expected subjective value *ESV*, which is not about future experienced values, but future expected values. In turn, subjective value *SV* is the result of a learning process through time, in which the subjective values adapt in $t + 1$. All subjective value relates closely with learning processes through time and with the mechanisms that generate expectations.

Reward predictions are not about realized *SV*s, but about expected subjective values, which makes the interactions recursive. So, the degree of my reward prediction error about actual subjective values depends on the expectations I had formed about my future predictions. At each moment of time, I anticipate what my expected subjective values will be, and these values are discounted and summed up to arrive at my current *ESV*. In contrast, experienced subjective value *ExperSV* is an abstract magnitude that relates certain quantities of a good x_j with a state of the neuronal system, that is, with the rates of firing of particular groups of neurons. This, however, is not taken as an absolute value but is corrected via the reference to a baseline value *b* which is a reference point, such as in prospect theory. This importance of context is further highlighted when recognizing the state of the organism (e.g., the degree of actual vitamin deficiency relative to the situation of eating apples). The *ESV* then results from the difference between the *SV* which had been predicted and the *SV* which was experienced. This prediction of *SV* involves the prediction of future reward prediction errors, which in turn relate to expectations of future expected values.

Glimcher recognizes that experienced subjective value is determined in brain areas different to those of expected subjective value. The same applies to general subjective value, which also implies a neuroanatomical differentiation between a decision-making system and an evaluative system, thus suggesting a molar-level dual selves approach. This is obvious if we consider malfunctioning, which can be most simply defined as systematic errors in the prediction of future states of the organism. In Glimcher's model, this is possible, as the predicted values and the experienced values are generated in different subsystems of the brain. Ultimately, the notion of individual welfare or well-being needs to refer to *ExperSV*.

So it seems that the discussion sometimes confuses the question of the consistency of the decision (sub)system with the question of the consistency between the Wanting and the Liking system. Even if the Acting Self was modeled as an internally consistent system, this would not preclude conflicts occurring between the Wanting and the Liking systems. In particular, observed behavior is a consequence of individual histories of learning how to balance the two systems and is most obvious in the wide variation between mixes of seemingly rational and irrational behaviors in every human population. This is especially the case with regard to dysfunctions such as addiction. There is no contradiction between the assumption that the decisions of the addicted person are "rational" and the dysfunctionality of realized behavior. This is precisely because the Wanting system is partly autonomous and could therefore be modeled according to the utility function approach if those utilities refer to subjective value as being different from experienced subjective value.

Conceptual Affinities With the Somatic Marker Hypothesis

I summarize this complex discussion in the following conclusion. The dual selves model which distinguishes between Acting Self and Object Self can be justified on principled information theoretic grounds. It is the most simple axiomatic alternative to unified models such as the economic concept of a utility function. Other versions of dual selves models can be subsumed under this most general model. This is especially true for all models that distinguish between different modes of choice, such as "planning" versus "doing," "planning" versus "habits," or "conscious processes" versus "automatic processes" (Camerer et al., 2005; Redish et al., 2008; Thaler & Shefrin, 1981). These can be seen as further specifications of the Acting Self and hence elements of a more complicated modular theory. At the same time, the model can integrate very general theories about choice which also present serious (though controversial) alternatives to integrationist and reductionist approaches to neuroeconomics.

This holds especially true for the somatic marker hypothesis (Damasio, 1995; Bechara & Damasio, 2005), which also encompasses more specific neuroeconomic models of reward and anticipation (Reimann & Bechara, 2010). In this approach, the dual selves model emerges as a distinction between brain states and bodily states, such that the Acting Self/Object Self distinction at least partly converges with the brain/body distinction. Somatic markers are bodily states that represent valuations of events based on past experiences, whereas the process of choice can be modeled in terms of, for example, the now standard model of reward prediction errors (Bechara & Damasio, 2005). To emphasize this again, the dual selves model does not take any empirical specification of functional differentiation of brain structure as a given. Thus, if the somatic marker hypothesis distinguishes between two different loops of valuation (i.e., the "body loop" and the "as if body loop"), this is exactly what can be expected, as the Object Self function can be manifested in structures as different as certain bodily states and/or certain states of parts of the brain. Via these loops, the entire decision-making process can also achieve consistent states via higher-order coordination mechanisms, in particular the emotions. Evolved emotions as anchors for choice and behavior are also seen as integrating forces in evolutionary psychology, which at the same time emphasizes the modularity of the brain (Tooby & Cosmides, 2005).

Having established the dual selves approach as one of the two cornerstones of my externalist approach to neuroeconomics, I now turn to biosemiotics. The connection between the two emerges from the observation that the Acting Self essentially relies on representations, that is, signs, to implement choices and decisions. Interestingly, this view is already presaged in somatic marker theory, but this approach locates markers exclusively within the brain-body system and so remains internalist. In the biosemiotic view, the internal signs qua markers would be essentially connected with external signs, thus relating somatic markets with evolved functions.

III. Biosemiotics: The "Bare Bones" Model

The fundamental idea in biosemiotics is that in all living systems the response by a system to a physical interaction with the environment is determined by *two* causal processes. These are the direct physical effect and the effect mediated through a sign. This duality results from the fact that living systems have an internal information structure in which the physical interaction is categorized according to certain criteria which have evolved through natural selection in the past. The physical interaction is *interpreted* in a particular way. Both the direct physical effect and the interpretation causally interact in generating the response of a living system, which in turn is subject to selective forces in the environment. To take a standard example, a frog's eye receives photons that manifest a particular pattern, and this pattern is interpreted as a "fly," triggering the response to catch it. The pattern as such may be that of a dark, small and moving object. This pattern is the sign of the object. The object itself is a physical composition of chemical substrates, some of which are nutrients relative to the frog which cause certain effects once digested. Obviously, the sign is not identical to these physical facts.

Peirce's Semiotic Triad, Naturalized

In this general characterization, I present a special version of biosemiotics which is strictly naturalistic. In this it is compatible with teleosemantics and related positions in the philosophy of mind. I also strictly follow Peirce's theory of signs which proceeds from a fundamental distinction between an object, a sign, and an interpretant, mainly following two recent contributions on Peirce: Stone (2007) and Robinson and Southgate (2010). This approach interprets meaning as reflecting evolved functions (on the discussion in semiotics, see Emmeche, 2002; Herrmann-Pillath, 2010; Vehkavaara, 2002). Functions refer to both the internal mechanisms that underly the effects of signs on an organism and the external embeddedness of the sign use in the larger evolutionary context. Insofar as functions emerge from an evolutionary process, the naturalistic biosemiotics approach is strictly externalist.

This is especially important when considering human agents for which mental categories are often taken for granted and to represent purely internal facts. So, for example, <snake> is a complex of perceptual patterns with snake-like shapes that establishes a sign which has direct effects on an organism that is confronted

with the sign and establishes a connection with the designated object or physical entity SNAKE (fear, arousal, tendency to flee etc.). These are internal causal mechanisms. Yet, the meaning of <snake> cannot be just reduced to those internal mechanisms, because the sign has complex and fuzzy boundaries, so that its meaning and function can only be explained in the larger evolutionary context. This context refers to the phylogenetic and cultural processes that drove the emergence of both the sign and the corresponding neuronal responses. Thus, meaning is completely reduced to function, both in the sense of meaning evolving out of selection (modeled e.g., in terms of evolutionary game theory, as in Skyrms, 2004, 2010) and in the sense of meaning supervening on functions (Adams & Aizawa, 2010; Macdonald & Papineau, 2006; Millikan, 1998).

In a first approximation we can describe the naturalistic Peircian framework as in Figure 1(a), which builds on similar diagrammatic expositions in the semiotics literature (e.g., Brier, 2008). I emphasize that this is a conceptual or metatheoretical model, and does not present any specific hypotheses relevant when analyzing the different pairs of interactions within the semiotic triad. Semiosis is triadic in the sense that the relationship between object and interpretant is twofold, direct, and intermediated via the sign. The SNAKE as a physical object is the source of a physical impact on the receptors of the eyes. But this is not sufficient to establish the presence of a "snake" in cognitive terms, because there are many things that may evoke similar physical effects on the eyes. Therefore, the reaction by the organism is mediated via the sign <snake>, which is actually a rule based definition of what physical impact will count as a "snake" cognitively as it refers to a SNAKE in the physical environment. This difference is central for the possibility of dysfunctions, or, in teleosemantic terms, misrepresentations. Therefore, the relation between sign and object is a relation of inference. Semiosis enables the organism to infer certain properties about an object which allow it to recognize the object as a SNAKE through perceiving <snake>. So, semiosis is an information generating and processing phenomenon (labeled inference in Figure 1a).

The information-generating role of signs emerges from the fact that semiosis is subject to selection in the evolutionary context. Selection impacts the final element of the triad, the interpretant. The interpretant is the element which is the vehicle of the inference process. It is absolutely essential to recognize that this is not a

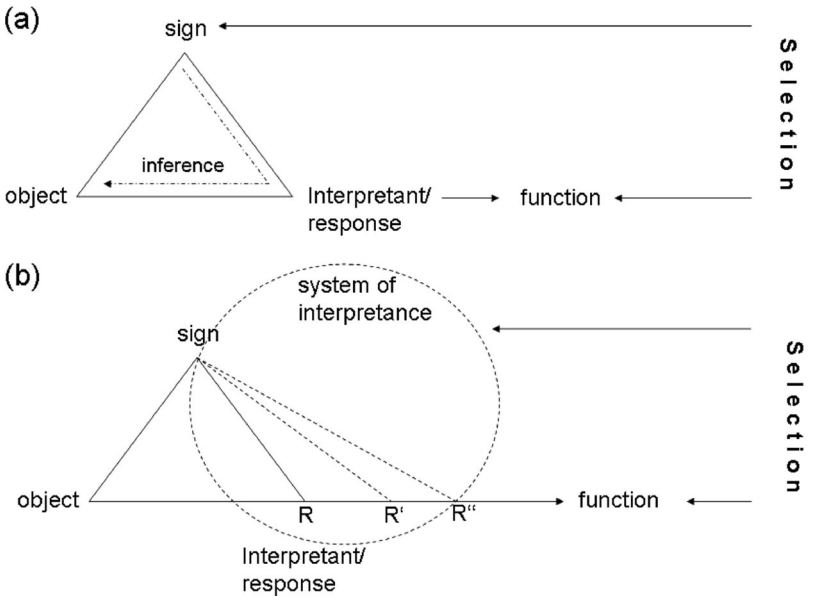

Figure 1. Basic structure of semiosis.

mental entity that gives meaning to the sign, but the reaction of a certain entity, such as the organism, to the sign. So, Stone (2007) calls the interpretant a "response" *R*. The response to perceiving <snake> is to flee. This implies that the information is actually carried in the response in the sense that the response reflects those properties of SNAKES that determined the evolution of the sign <snake>, that is, snakes might be dangerous to survival. Both the interpretant and the sign have evolved under natural selection, that is, we can speak of a coevolution of the two sides of the semiotic triad. The information content of the response consists in the realization of a function, in our example, to ensure survival.

This general structure of semiosis can be analyzed in more detail by introducing the idea of a "system of interpretance" (Salthe, 2009; Figure 1b). This refers to the hierarchical ordering of functions. The relationship between flight and survival is essentially generic. There are many intermediate functions which relate to the complex structure of an organism, including the ecological context. So there is a chain of responses *R*, *R′*, *R″* and so forth, which are part of such a system. The important point that results from this perspective is that semiosis has the property of multiple realizability. This means, the same generic function can be realized in different forms. For example, one can be flight, or another stamping on the snake. Robinson and Southgate (2010), following Stone (2007), therefore introduce the distinction between the response *R* and the purpose *P*, with the latter being the generic aspect under which selection takes place, in the sense of selecting *for* (fulfilling a purpose, as opposed to selection *of* the traits that fulfill the purpose).

The Complete Biosemiotic Model

Further following Robinson and Southgate, we can now proceed to draw a more detailed picture of the semiotic triad, which is central for establishing the relationship to neuroeconomics (see Figure 2). This makes the distinction between the semiotic mode and the physical mode more explicit. Using the notation in Robinson and Southgate (2010), but modifying and extending their model, I introduce a twofold triad (see Herrmann-Pillath and Salthe, 2011; see Figure 2).

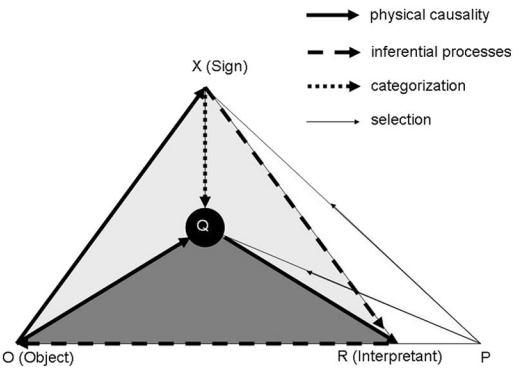

Figure 2. The extended semiotic triad.

The inner triad *O-Q-R* is the physical chain that undergirds the semiotic relation, such as the physical impact of an object on the eyes and the subsequent chain of neurophysiological effects. This relationship is accompanied by the sign relation *O-X*, also physical, as the sign is caused by the object. However, this second relationship can be complex. This underlies Peirce's fully fledged theory of signs which I do not wish to consider in this paper. For example, the relation can be directly caused, such as when fire causes smoke (an index), or it can be arbitrary, as in the case when a sign is a symbol. However, even in the latter case, the sign remains a physical phenomenon. This discussion shows, however, that the sign relation, though a physical one, differs fundamentally from the physical relation *O-Q-R* because strictly speaking, the underlying causal process moves in the opposite direction. This is easy to see when we ask the question *why* the sign is an effect of the object (Stone, 2007). A full answer must refer to the entire semiotic process, claiming that it has been assigned to the object via an evolutionary process. In contrast, the relation *O-Q-R* is not directly dependent on the evolutionary process: A SNAKE will impact the receptor's eye regardless of whether it is seen as a "snake" via the intermediation of <snake>.

The relation *X-R* is also physical of course, as the perception of the sign is a physical process. This is not shown here using a continuous line because this relation is conceptually dominated by the inference process. These are indicated by the broken arrows, also showing the inverse causality. The sign reveals information about an

object, such as when perceiving <snake> reveals that an object is a SNAKE. It is precisely for this capacity that selection has resulted in the emergence of the sign. We can then say that the semiotic relation *O-X-R* supervenes on the physical relation *O-Q-R*.

This brief summary of the biosemiotic approach suffices at this place. My aim is to show now that the duality of the physical and the semiotic causal channels directly reflects the analytical distinction between molecural and molar levels of analysis in neuroeconomics. In doing this, I argue that this is not merely a methodological distinction, but reflects some fundamental properties of neuronal systems, seen as products of an evolutionary process.

IV. Semiosis in Dual Selves

I distinguish between the organism as a physical entity and the interface with the physical environment on the one hand, and the duality of Object Self and Acting Self on the other. The central conceptual link between the two results from the focus on the generating and processing of information. Semiosis is an inferential process and for that reason complements the dual selves structure as this features essential incompleteness of information, in particular that internal to the physical boundaries of the brain. This is the distributed cognition perspective: Externalized cognitive mechanisms are essential to brain functioning because of the necessary incompleteness of internal mechanisms.

In the notion of the interface, external world and organism are conflated in the sense that there are no other sources of information about the external world apart from sensory data. This has the important implication for economics and the theory of consumer choice that a good is always a set of sensory inputs and cannot be naively equated with the physical entity (which, however, fits into the long-standing Lancasterian tradition of modeling goods in certain economic models). At the interface, information is generated about that physical entity. Therefore, in the triadic diagrams we show two broken arrows leading back toward the object, and representing the information-generating dynamics of the evolutionary process. This information is not representational, however, but relative to the ultimate criteria of selection which drive the responses triggered by the evolving information.

The Object Self then corresponds to the response *R* in the semiotic approach in so far as it is the instance of the evaluation (i.e., interpretation) of the outcome of an action, as response *R*, that has been generated by a physical chain of causation *Q* which connects external world and outcome. Other than the standard economic approach, this implies that the interpretation, that is, response *R*, relates to a general purpose, that is, a function that was selected for multilevel evolutionary process in the past. Therefore considering current economic research, this part of the picture can be grasped by evolutionary approaches to preferences (as proposed e.g., by Robson & Samuelson, 2010; for consumer theory, see Saad, 2007). However, in the dual selves model these preferences *do not* direct choices. Choices work via the semiotic intermediation, which relates to the Acting Self. This fits the general characterization of the Acting Self being a cognitive agency which operates on the basis of representations of the external and internal world, that is, signs. We can, therefore, locate the dual selves approach in the triadic framework of semiosis (see Figure 3).

The relation between the Acting Self and the Object Self is mediated via the semiotic categorization of physical chains *Q* that link organisms and evaluations. For example, the organism is in need of nutrients, and the Object Self evaluates certain intakes according to certain criteria, such as the most simple, distinguishing between hunger and satiation. However, the selection of those actions that lead to food intake

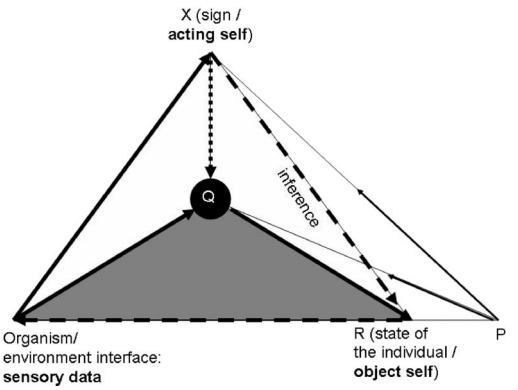

Figure 3. Object self and acting self in the semiotic triad.

works via the Acting Self. The Acting Self categorizes the different action patterns according to their semiotic intermediation. This implies that the relation between Acting Self and Object Self is open and indeterminate. For example, choice among food items follows indirect features such as color or taste. These do not directly reflect nutrition values. In fact, the Acting Self is ignorant about the physical causality between food intake and states of the Object Self. These chains are actually the object of biological and medicinal inquiry, knowledge of both of which is still fallible today. This picture corresponds to recent approaches in neuroeconomics which argue that preferences are constructed via consumption activities and integrate environmental factors via memory (Ariely & Norton, 2007). In this case, the Acting Self learns from the observed functionings of the Object Self in terms of their external embeddedness. In short, preferences *are* the actions qua responses *R,* which are constructed by the Acting Self by means of semiotically intermediated proprioperceptions. This is the externalist neuroeconomic complement to the notion of revealed preferences in the standard economic approach: *There is no such thing as a mental state that is a "preference."*

This analysis has far-reaching consequences for the conceptualization of human choice. On the one hand, the dual selves approach started out from stating an internal information asymmetry in the brain. I have shown that this internal asymmetry directly matches the triadic analysis of semiosis. What, however, does this imply about the nature of the information that is processed by the brain?

First, the Acting Self reflects the semiotic level as it is a decision making module that processes signs. Although this also involves sign processing internal to the brain, the most important signs are external to the brain. In the context of human societies, many signs are culturally determined and transmitted. In this sense, the Acting Self is a public entity as it operates on the basis of population level sign processes (this is reflected in the strong impact of social interactions on individual agency, of which, however, the human agent is to a large degree not aware, see Oullier & Kelso, 2009; Wegner & Sparrow, 2007). Second, the Object Self is by no means a purely internal phenomenon, but relates to functions in the larger context. Therefore, even though we might locate experiences in the individual, in fact what is behind them is an evolved pattern of internalized functions in a larger context. This viewpoint has been elaborated on, both in the context of general evolutionary theory and in the theory of consumer behavior. In other words, the preferences of the Object Self are not simply irreducibly subjective but relate to evolved functionings of the individual, themselves objective in terms of a possible evolutionary analysis. This refers to both the biological and the cultural or societal level (as outlined e.g., by Jablonka & Lamb, 2006). Thus, both the Acting and the Object Self operate on the basis of information that is "environmental information" in the sense of information theory, that is, that does not relate to meanings involving the intentions of senders and receivers (Floridi, 2003; Aunger, 2002; Skyrms, 2010). Both signs and evolved responses are population-level phenomena and hence external to the individual, though essentially relying on neuronal mechanisms (which are physically "internal" to the body/environment interface) for their actualization.

We have now all the essential analytical tools at hand to discuss the question of how far neuroeconomic externalism is more powerful than neuroeconomic internalism for exploring the relation between molar and molecural processes, that is, theoretically enriched observations about observed behavior, such as behavioral data, and the underlying neuronal mechanisms. Consumption disorders are the acid test, as they are the most important falsifiers of the standard economic model which would need to be accounted for by any attempt at a neuroscience reduction. I will begin with a methodological case study of one influential article in economics, Laibson (2002).

V. Merging Neuroeconomics and Semiotics in the Analysis of Consumption Disorders: The Example of Smoking

I am going to present a semiotic reconstruction of what Laibson (2001) has proposed as a "cue theory of consumption" in the context of economics, which builds on a long tradition in psychological research, and takes up important aspects of existing psychological and neuroscientific approaches to addiction. I have chosen the Laibson model as a case study because it maintains the formal charac-

teristics of the standard economic model as closely as possible and so clearly fits in with the formal structure of Glimcher's approach. On the other hand, the notion of "cues" directly connects with the semiotic approach, as cues are signs. I, therefore, take the economic model proposed by Laibson, which on first sight seems to integrate easily with Glimcher's approach, as a point of departure for developing a strictly externalist approach to choice, which stands in opposition to reductionist neuro-internalism.

Cues and choice: The Economic View

Laibson has presented an extension and modification of the original Beckerian approach to habit-formation, including the so-called theory of rational addiction (Becker & Murphy, 1988; for related theories, see, e.g., Bernheim & Rangel, 2004). The central idea is that environmental cues elicit changes in the preferences of individuals, thus inducing shifts in the marginal utilities of consumption. This happens because cues activate memories of past consumption. Our example will be smoking, so, for example, perceiving a box of cigarettes activates a strong desire to smoke, because the cue moves the schedule of expected marginal utility as well as the actual marginal utility of lighting the cigarette. This cue-theory of smoking corresponds to a long series of results in neuroscientific research which have highlighted the role of environmental factors in driving smoking behavior as compared to the purely pharmacological factors (e.g., Caggiula et al., 2002; Hughes, 1989; Le Folla and Goldberg, 2005; Reid, Ho, & Berger, 1996).

What is a cue? First, a cue is a sensory input. But this sensory input is different from the sensory input that actually creates the utility of a particular good and activity. There is a spectrum of possibilities here, ranging from arbitrary environmental items to the physical embodiment of the respective good. So, for example, if somebody mostly smokes in a pub, her craving for nicotine will differ depending on whether she is in a pub environment or a gym. Moreover, and most interestingly for our discussion, the cigarette itself is a cue. This differs from the physiological chain connecting the activity of smoking with nicotine infusion into the organism. This includes proprioperceptive cues, such as the scratch of the smoke in the throat (West, 2006). Thus, if the different properties of the cue of the cigarette are changed, the marginal utility of smoking will change, too. So, biologically, cues relate with Pavlovian reinforcement mechanisms which change the individual's dispositions to react to a stimulus, both in the sense of preparing the organism and also in that of compensating for negative effects.

The Laibson model grasps these effects by introducing a metautility function which includes the cue-based physiological mechanisms seen as biologically determined. The standard utility function is thus embedded into this metafunction in the sense that the actual choices of the individual become dependent on the history of consumption experience. These have evolved as a result of certain patterns of cues and correlated behaviors, and are, therefore, contingent on the environment (this establishes a relation with the human capital approach to consumption developed by Becker, 1996). An important consequence of this model is that individuals who are aware of the role of cues in eliciting behavior will possibly be willing to manage cues directly in order to influence their own behavior. So, for example, a smoker might avoid pubs if he intends to quit smoking, in order to avoid the leveraged cravings which would also cause greater pain of nonsmoking. This also implies that cues without subsequent consumption are hedonically aversive. The theory can also explain the phenomena which are normally associated with hyperbolic time preferences, that is, a heightened preference for present consumption. If, in the present, a cue impacts on choices, momentous marginal utility is leveraged, whereas the anticipated utility of future consumption streams is only based on cue-less evaluations, since the cues must be physically present in order to operate. A final core insight of the cues-theory is that consumption is a public activity, because observed consumption of others is a cue itself, so that public consumption generates positive externalities in other individuals, whose marginal utilities are increased by the cues. Smoking in companionship increases its pleasures (on the impact of observed smoking, even imaginary as in movies, on smoking behavior, see Wagner et al., 2011).

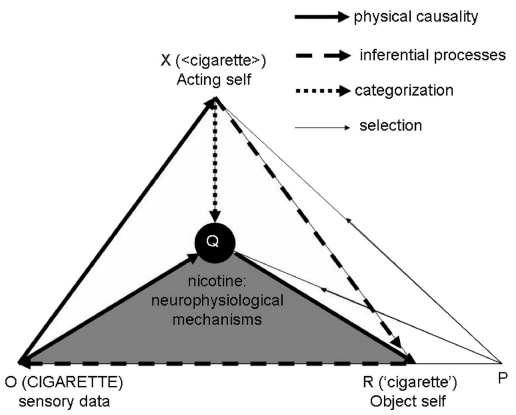

Figure 4. The biosemiotic view on cues in smoking.

Cues and Choice: The Biosemiotic View

Laibson's theory is a clear-cut example of a semiotic process as outlined in the previous sections (see Figure 4). Cues are signs. Putting the smoking example into the triadic framework, we immediately recognize that the cue qua sign mediates the *O-X-R* causal chain and thereby affects the functioning that is active in the direct chain *O-Q-R* (which is, economically, the schedule of marginal utility, in the views of other disciplines, neuropharmacological causality). The latter chain refers to the neurophysiological effects of nicotine in the body, causally triggered by the CIGARETTE, the former chain relates to the sign of the CIGARETTE, the <cigarette>, and what smokers smoke, are "cigarettes." Therefore, smoking has to be analyzed into two different processes, one is the inhalation of smoke, and the other is the perception of the cigarette—a sign. Both processes converge in the neurophysiological responses of the smoker, that is, the response which is the level of satisfaction generated by the activity of smoking. The "cigarette," relative to the Object Self, is not simply the physical object with a certain amount of nicotine. The sign changes the response *R* of the individual to the nicotine infusion by changing the degree of satisfaction, as compared to physically similar infusions of nicotine without activation of the sign (West, 2006). Hence, Laibson's notion of changing marginal utilities actually refers to an interpretive process. The original object, the CIGARETTE, triggers two different causal chains that coevolved during the process of habit formation. This process can be analyzed as resulting from an evolutionary dynamics on different levels, with the central ones being the long term selection of the underlying biological mechanisms, the medium term of the cultural evolution of artifacts and the short term of individual habit formation (on the applicability of the fundamental selection model on different levels, including individual reinforcement learning, see Hull, Langman, & Glenn, 2001; Mesoudi, Whiten, & Laland, 2006).

In terms of the semiotic triad, we can further analyze this process. The CIGARETTE is actually a set of sensory data, including those both consciously perceived (such as the touch on the lips) and those unconscious (such as the inhalation into the lungs). In this sense, one could even say that the physical object CIGARETTE perceived as such directly relates to the sign <cigarette>, and the object disintegrates into a large set of sensory data when actually put into use. These data are semiotically reconfigured as "cigarette." In Laibson's framework, emphasis is on the equilibrium states. Dynamically, the central process is governed by the coevolution between the *O-Q-R* and the *O-X-R* chain. This is reflected in the semiotic fields of the notion of a <cigarette>, that is, the shape, color, size, and so forth of what counts as a cigarette in a particular cultural setting and for a particular smoker. This relationship can be broken down to the brand level and hence also applies to the analysis of advertisements and individual consumption profiles. Dynamically this is also a learning process in which selection establishes increasingly sophisticated consumption patterns through individual learning. The semiotic approach allows assigning a systematic place to all these processes. They are the standard topics in psychology, consumer research and cultural studies, with different levels of generality, ranging from individual idiosyncrasies to larger cultural patterns. All of this relates to the level of the Acting Self, which actually deals with the signs in order to achieve certain states of the Object Self.

The basis for all of this is the neurophysiology of nicotine effects in the organism which is the physical foundation for the semiosis of smoking. However, as has been demonstrated in

the empirical data as reported by Laibson, this is not sufficient to describe the response *R*. Even if the Acting Self were perfectly informed about the neurophysiology, it would not be able to track optimal behavior in terms of states of the Object Self. These states depend essentially on the environment in which the action takes place. The Acting Self operates on the basis of the semiotic representation of the environment. Correspondingly, the states of satisfaction attained by the Object Self are themselves semiotically determined, which, however, reverts to a past emergence of functions related to the activity of smoking. These functions also go beyond the mere neuropharmacological effects in that they include broader contexts of the activity of smoking, such as its role in social communication.

The semiotic analysis makes plausible why in the explanation of smoking behavior there is no clear-cut relationship between the psychopharmacological properties of nicotine and smoking, even when single individuals, who might change their behavior over time, are taken as one unit. There is no unique mapping between the molar and the molecular level, hence no way to fully reduce the former to the latter. The observed response *R*, smoking, is determined by the complex interaction between neuropharmacological factors, their semiotic categorization, the individual information processing resulting from semiotically intermediated inferences and the functional embeddedness of the response. Given the complexity of this system, there are a great many possible developmental pathways which can be triggered even by very minor events (West, 2006).

To summarize, I argue that the Laibson model, which is a strictly economic model based on (almost) standard utility theory, in fact stands at a crossroads from which either neuro-internalist reductionism or the externalist approach of biosemiotics could be reached. In the former case, the notion of utility would be rendered in internalist terms, thus contradicting the externalism of the standard revealed preference approach to utility, which seems to result in a serious contradiction. In the next section, I show that the more general and highly diverse approaches to consumption disorders and addiction in different fields of psychology and cognitive sciences clearly justify the biosemiotic view. Only this view matches the externalism of the standard economic model with externalist neuroeconomics. At the same time, what is a strong falsifier to the standard economic model, turns into a confirmative exercise in applied biosemiotics.

VI. Generalizing the Externalist Argument: A Semiotic Framework for Hypotheses on Consumption Disorders and Addiction

Eating Signs: Biosemiotic Analysis of Consumption Disorders

In most contemporary developed countries, smoking is generally considered as a dysfunction, independent of its extent. Eating as such certainly cannot be counted as a dysfunction, only eating disorders can be. This means that food consumption is another excellent example for a generalization of the analysis above. Recently, economists have devoted substantial efforts to understanding eating behavior, not only because obesity is a major dysfunction in modern societies, but also because the mere availability of information about appropriate nutrition does not help to solve the problem (Downs, Loewenstein, & Wisdom, 2009). Instead, there is increasing evidence that engineering the cues that guide eating behavior is a more successful strategy. This has recently been dubbed "nudging" (Thaler & Sunstein, 2009; compare Wansink, Just, & Payne, 2009). I claim that "nudging" lends itself easily to a semiotic analysis.

The analysis of eating behavior also allows for more direct reference to the brain structures behind consumer choice. Food consumption is a multilevel process that involves many different kinds of organismic activities, reaching from closer-loop metabolic circuits to highest-level conscious choices. It is also a most fundamental human need, so there would be good reason to expect that, over the course of evolution, information processing about food consumption would have been optimized. However, precisely the split between Acting Self and Object Self can be seen as the precondition for the high degree of phenotypic flexibility and hence, increased adaptive performance in varying human environments. In this sense, growing internal information deficits simply reflect the increasing advantages of phenotypic flexibility (cf. Richerson & Boyd, 2005). As a result, individ-

ual learning also plays an essential role in the ontogenesis of eating habits.

The basic brain structure that underpins eating behavior reflects the dualism between the Wanting and the Liking systems, as surveyed by Berridge (2009) (see also e.g., Finlayson, Halford, King, & Blundell, 2010). On the one hand, food consumption is regulated by general homeostatic mechanisms which indicate a state of well-being that results from food intake, such as satiation. There are special circuits for indicating satiation and others for the hedonic response. This Liking system differs fundamentally from the Wanting system which treats food as one item in the generalized set of rewards and which focuses on the mesolimbic dopamine circuits. The two systems interact twofold. One is that the Wanting system is a separate source of rewards, as the perception of cues that indicate and accompany food raises the dopamine level. The other is that the Wanting system may change the sensitivity of the Liking system. This can be basically conceived as a Pavlovian mechanism of reinforcement learning, such that certain cues of food availability become the main drivers of the actual behavior.

In the case of food, the role of cues is most obvious from the role of visual and other additional sensory indicators such as taste accompanying food consumption (compare Simmons, Martin, & Barsalou, 2005). Thus, a dish that looks delicious can be a source of a double reward obtained from eating. First, the food intake activates the homeostatic mechanisms that signal well-being (the *O-Q-R* channel in the semiotic triad), but second, the Wanting system may be especially activated by the sensory aspects of the food as these trigger an independent dopaminergic response and hence reward. The reward in turn might change the homeostatic mechanism (the semiotically intermediated categorization of the *O-Q-R* channel). So, eating delicious-looking food may result in eating more than is required just for satiation from the perspective of the Object Self. Interestingly, there is also an independent effect of the signs on actualized eating behavior. Imagining eating food later reduces the actual quantity consumed, presumably because the habituation to the sign input simultaneously changes the sensitivity to the object input (Morewedge, Huh, & Vosgerau, 2010).

The testing ground for this dual selves approach are the dysfunctions of the system. In the integrated approach, which projects into the economic notion of utility, the distinction between different standards of evaluation does not matter because in the end the nutritional effects and the effects resulting from aesthetics are both rewards in terms of utilities. In the semiotic framework, actual behavior results from the confluence of different causal forces, with the Acting Self operating partly independently from the evaluations of the Object Self as a result of the underlying fundamental information asymmetries. The underlying dopaminergic circuits can then become partly decoupled from the signaling of the homeostatic mechanisms, and hence an independent source of satisfaction. In this perspective, eating behavior is partly transformed into the pure consumption of signs, underpinned by a peculiar selection of mechanisms *Q*. Indeed, eating disorders are an excellent example of this distinction between the consumption of signs of goods as compared to the consumption of goods. Compulsory eating of sweets does not result in the fulfillment of nutritional functions, but can even end up in organismic dysfunctions, and in many cases also in less perceived well-being. So it can involve a stark contrast between experience and decision utility, as frequently the individual feels very bad after the action and even tries to restore the original state by compulsive action such as vomiting. In this way, the choices of the Acting Self are virtually decoupled from the Object Self. This implies the distinction between the "good" and the "sign of the good." In the compulsory eating of sweets, the individual actually consumes the signs of goods, the <sweets>, and the accompanying temporary feeling of satisfaction results from the activation of the underlying dopaminergic circuits. The SWEETS, however, damage the organism, and lead even to perceived dissatisfaction on the part of the Object Self.

This analysis corresponds to related applications of the model of dopaminergic circuits which are triggered by cues of reward. For example, in gambling, the feeling of pleasure is caused by the fact that in most gambles, the impression is very often created that a player has only just missed the target by a small deviation (for a recent survey of research on gambling, see Clark, 2010). This small deviation is a cue that indicates rewards. This reinforces the gambling activity, possibly to a degree that the

constant pursuit of a target which is actually mostly missed creates an independent source of satisfaction resulting from the increased levels of dopamine. In other, almost paradoxical words, a gambling addict does not enjoy the gains, but the recurrent small deviations from gains. The Acting Self relentlessly pursues the signs of gains, but the Object Self suffers ever increasing losses. This analysis matches with the claim made by Ross et al. (2008) that gambling is the prototypical case of addiction, as it does not involve any direct psychopharmacological causalities, but only dysfunctions in information processing. Indeed, there are good reasons to speculate that the semiotic duality of causal pathways may relate with the interaction between dopaminergic circuits and glutamate processing circuits, as the latter sensitize the brain to certain sensory inputs which are predictors of rewards, that is, signs in my parlance (Ross et al., 2008). This interaction is generic and independent from specific neuropharmacological determinants of addiction to certain substances.

The semiotic approach also makes plausible why there are fuzzy boundaries between consumption idiosyncrasies (such as collector's manias), consumption disorders and addiction. Ascoli and McCabe (2006) argue that all scarce goods can become the object of addiction in this sense, given the necessarily strong involvement of dopaminergic circuits in aiming at obtaining scarce goods. Lea and Webley (2006) distinguish between pharmacological and perceptual drugs, with the latter including gambling or pornography, but even also money, under certain circumstances. Although this extension of the notion of addiction does not represent majority opinion in the pertinent literature, the rationale of these arguments lies, in my analysis, in the complex interaction between the two triadic channels. Depending on their complex interaction, there is the possibility that, in individual cases, a constellation of neurophysiological factors, external functionings and semiotic intermediations emerges and this results in the transformation of consumption idiosyncrasies into quasi-addictive behavior.

It is important to notice that the signs themselves are not human universals. This is why we need fully fledged semiotic analysis to fully understand human eating habits. The view of chicken feet may trigger craving for food in the Chinese, but may appear to be repulsive for most Western Europeans. Eating behavior is based on customary behavior, the basis of which is partly externalized in the material environment. The environment is actually a complex repository of implicit information that guides consumer choice. For example, Sobal and Wansink (2007) have presented a detailed analysis of eating microenvironments, labeled "Kitchenscapes," "Tablescapes," "Platescapes," and "Foodscapes" which guide eating behavior. Cultural conditioning of eating does not mean that individuals have internalized values of preferences, but is in fact a coevolution of environmental features and a generalized reward processing mechanism in a dual-selves structure.

Locating Hypotheses About Addiction in the Biosemiotic Framework

As an effort to systematize the previous considerations, I propose a generalization of my argument, based on two comprehensive surveys of addiction research (Redish et al., 2008; West, 2006). The theory of addiction is a workhorse for dual selves approaches, and it achieves a central role in the theory of consumer choice if we consider the fact that, in recent theorizing about addiction, the most comprehensive approaches combine different kinds of mechanisms and the different levels on which those work. Basically, as in West (2006), those theories relate certain inputs (drugs and drug-related sensory inputs) with certain behavioral outputs. The inputs are processed on different levels of complexity, ranging from automatic responses such as impulses through motives to evaluations and plans. For these, certain background forces are essential which relate to generic drives and emotional states. These models also include mechanisms of self-control which link up with concepts of identity in social context. If we strip them down to a basic and more simplified conceptual structure, we can distinguish between two fundamental aspects:

• The first is the distinction between direct neurophysiological causal pathways and semiotically intermediated ones. The former includes all the direct and indirect ways by which the intake of certain substances and the repetition of certain actions directly change the neurophysiological mechanisms that drive decision-making, similar to the way drugs change the levels of dopamine in the respective brain

modules and hence change perceived rewards, different forms of neurotoxicity of drugs, or, possibly most importantly, the opponent process theories (West, 2006). I also include all opponent processes that involve compulsive actions under this heading, because they result in truncated action sequences where the expected outcome of the action does not matter at all for the triggering of the action (the action is the goal) (Redish et al., 2008). In contrast to this, the semiotically intermediated pathways include all the phenomena of cue-based alterations of choices (such as incentive-salience models; West, 2006), as well as all higher-level processes which involve the reflection of choices, such as self-control based on self-perceptions. The latter always involve reference to the symbolic representation of states and actions, such as frames determining the estimation of probabilities or social categories involved in assigning identities to agents.

• The second is the distinction between Acting Self and Object Self, which partly cuts through the common conceptual divisions in psychology and economics. This is because some aspects of addictive behavior refer to the Object Self directly, with a continuum of relevant phenomena. On the one hand, we have all the forms of automaticized actions which do not directly involve choice, or where choice is also driven by the direct reward of repeating the action pattern as such, independently of the reward resulting from the action (such as automatically lighting the cigarette without even considering the hedonic effect). On the other hand, this includes all the higher-level states of the Object Self which directly reflect the functions relative to the environment. These include self-conceptions and even plans which are directly related to certain aspects of the environment and in fact represent externally provided action programs (such as making a plan to stop gambling in order to be restored to the proper role as the head of the family). In contrast, the Acting Self involves all phenomena which affect the choice mechanism. This boils down to the basic dopaminergic circuits, but allows for a great deal of leeway either via the direct recategorization of neurophysiological pathways (such as leveraging the reward via cue-based adaptations) or via the altered cognitive representation of choices (such as changing perceptions of future outcomes in terms of probabilities).

I summarize this in Figure 5. In this figure, the location of the semiotic triangle in the center has an explicit meaning in the sense that the different parts of the triadic diagram are located in the boxes of the table in which they count most.

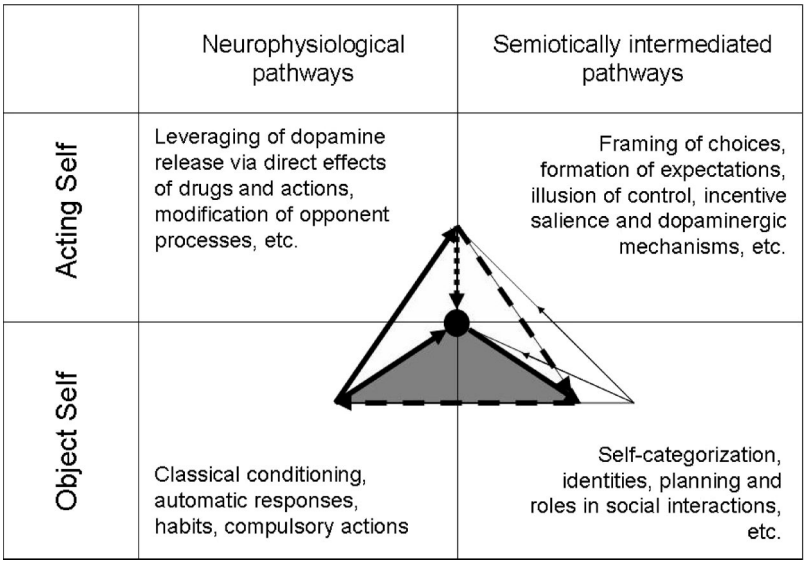

Figure 5. Semiotic analysis of addiction.

In the upper left-hand box we find mechanisms in which the conjunction of an object and the sign directly recategorize neurophysiological paths *O-Q-R*. An example is the changing dynamics of opponent processes during addiction (West, 2006). This is the direct effect of physiological dependency on a drug, which results, for example, from the use of the drug not mainly to create hedonic reward, but to avoid the effects of withdrawal. These effects are independent from semiotic recategorizations and, therefore, directly work via the *O-Q-R* channel.

In the upper right-hand box, we find all the processes which change the perceptions of actions taken, such as the illusion of control phenomenon, or the simple perception of costs and benefits. This affects the process of inferences which conveys information about the object, the drug, to the Acting Self, such as in the case of incentive salience, when the cognitive system is geared to overemphasize cues which relate to the drug.

In the lower left-hand box we find all the mechanisms by which the Object Self is transformed so that the action-goal relation is suspended, that is, the action becomes the goal, such as in habit formation.

Finally, in the lower right-hand box we find all the semiotically intermediated mechanisms which relate the Object Self to functions in a certain social and cultural context. This includes the self-categorization of individuals, activating a particular semantic field related to perceptions of social roles, or the making of plans, which always build on preconceived notions of the contextual determinants of plans.

VII. Conclusion: Externalism and The Cross-Disciplinary Integration of Theories About Human Choice

The gist of the argument presented so far is the claim that a proper conceptual framework for neuroeconomic research has to be an externalist one. As such, externalism stands in direct contradiction to Glimcher's (2011) model of internalist neuroscience. However, the example of addiction research clearly shows that this externalist view differs substantially from previously held externalist views in psychology and related disciplines. These focus on social and cultural factors in determining addiction phenomena (e.g., Griffiths, 2008; Lende, 2008). *Externalist neuroeconomics* puts neuronal processes and structures at the center of analysis. But these mechanisms are seen as being incomplete in a nontrivial way, which blocks any attempt at full-scale reductionism. I argue that this approach can be most generally characterized as a distributed cognition or extended mind approach and I offer a specification of this approach in terms of biosemiotics, meaning that I regard signs as the essential medium of externalist causalities in producing behavior.

Another difference between this view and the standard conjunction of neuroeconomics and behavioral economics is that the standard of malfunctioning is fundamentally different. Biosemiotics involves reference to proper functioning in contexts of selection, whereas the current neuroeconomic research mostly identifies anomalies in terms of reference to the standard economic model of rational decision making. However, the standard economic model fails to identify empirically valid criteria of malfunctioning. This is demonstrated by the example of addiction, the experimentum crucis here. The empirical record can be stretched to a considerable extent but the celebrated Becker and Murphy (1988) model ultimately fails to cover phenomena such as the context-dependency of cravings or the strong role of classical conditioning in the development of addictive behavior. It does not address one fundamental aspect of addiction, which is the perceived and experienced absence of choice at the moment when the behavior is elicited. And most importantly, it cannot explain the problems of recurrent relapses, which take place despite a long history of learning about the negative consequences of addiction; individuals accumulate experiences with painful withdrawals after rational choices to put an end to their behavior, only to relapse later again, with many repetitions (Ross et al., 2008).

On the other hand, the standard model cannot grasp the fact that behavior which is dysfunctional in terms of the model can be functional in relation to the context of actions. This raises the fundamental question whether the concept of "anomalies," which drives much behavioral economics research, is valid at all. Even an addiction can be functional relative to a certain context, which explains why addictive behavior can be suppressed in different environments, but relapses occur frequently once the original

context is reestablished. These functionings, in the biosemiotic approach, cannot, however, be explained as reflecting optimal choices. They emerge from the interaction between semiotic and neurophysiological mechanisms which produce organismic responses, that is, behavior which is reinforced in certain selective contexts, such as the prevalence of certain habits in a subpopulation. The concept of anomalies overlooks the fundamental fact that we cannot assume perfectness of information within the individual. This applies to both the relation between brain and body and to the relation between different modules of the brain. One driving force of patterns of distributed cognition is to produce functions in which the internal information deficiencies are resolved by externalized cognitive mechanisms. These mechanisms range from physical entities such as objects and physical arrangements of things, including both natural and artificial phenomena, to social cognition, that is, the leveraging of individual cognition by social interaction, such as contagion in social networks. Thus, in the case of addiction, the dual selves *cum* biosemiotics approach can address the fact that all phenomena of addiction also have a population level aspect, that is, individual behavior is strongly influenced by patterns in the social environment.

The proposed approach is in agreement with the incipient attempts at establishing a social neuroeconomics to especially highlight the coordination of behavior across different agents (Oullier et al., 2008, 2009). This shift of emphasis also reflects the philosophical approaches to agency which emphasize the role of language and narrative structures in social contexts (Ross, 2007). Most generally, it comes close to what has been called an "ecological approach" to rationality by Vernon Smith (2003). This investigates the contextual factors impacting on consumer choice, using the distinctions between different levels (for a related view in consumer research, see Foxall, 2008). Ecological research is multidisciplinary, involving both the sciences and the humanities. The sciences are invoked when certain patterns of organismic adaptation are under scrutiny. The humanities are necessary to identify the meaning of signs in particular contexts of social interaction, such as culturally contingent food habits. The same phenomena can be seen from different angles in order to achieve a full understanding of the underlying semiosis. So, the argument presented in this article offers a synthesis of different methodologies in understanding human behavior, while at the same time recognizing the pivotal role of the brain sciences for the future research agenda in this field.

References

Adams, F., & Aizawa, K. (2010). Causal theories of mental content. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2010 ed.). Retrieved from http://plato.stanford.edu/archives/spr2010/entries/content-causal/

Ariely, D., & Norton, M. I. (2007). How actions create—not just reveal—preferences. *Trends in Cognitive Science, 12,* 13–16. doi:10.1016/j.tics.2007.10.008

Ascoli, G. A., & McCabe, K. A. (2006). Scarcity begets addiction (Commentary on Lea and Webley), *Behavioral and the Brain Sciences, 29,* 178. doi:10.1017/S0140525X06249045

Atkin, A. (2009). Peirce's Theory of Signs. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition). Retrieved from http://plato.stanford.edu/archives/spr2009/entries/peircesemiotics/

Aunger, R. (2002). *The electric meme. A new theory of how we think.* New York: Free Press.

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior, 52,* 336–372. doi:10.1016/j.geb.2004.06.010

Becker, G. S. (1996). *Accounting for tastes*. Cambridge, MA: Harvard University Press.

Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics, 117,* 871–915. doi:10.1162/003355302760193913

Bernheim, B. D. (2009). The psychology and neurobiology of judgment and decision making: What's in it for economists? In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 115–126). Amsterdam: Elsevier.

Bernheim, B. D., & Rangel, A. (2004). Addiction and cue-triggered decision processes. *American Economic Review, 94,* 1558–1590. doi:10.1257/0002828043052222

Berridge, K. C. (2009). "Liking" and "wanting" food rewards: Brain substrates and roles in eating disorders, *Physiology & Behavior, 97,* 537–550. doi:10.1016/j.physbeh.2009.02.044

Brier, S. (2008). *Cybersemiotics. Why Information Is Not Enough!* Toronto, Canada: University of Toronto Press.

Brocas, I., & Carrillo, J. D. (2008). The Brain as a hierarchical organization. *American Economic Review, 98,* 1312–1347. doi:10.1257/aer.98.4.1312

Burch, R., Charles Sanders Peirce. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2010 ed.). Retrieved from http://plato.stanford.edu/archives/spr2010/entries/peirce/

Caggiula, A. R., Donny, E. C., White, A. R., Chaudhri, N., Booth, S., Gharib, M. A., . . . Sved, A. F. (2002). Importance of nonpharmacological factors in nicotine self-administration. *Physiology & Behavior, 77,* 683–687. doi:10.1016/S0031-9384(02)00918-6

Calvin, W. H. (1998). Competing for consciousness. A Darwinian mechanism as an appropriate level of explanation. *Journal of Consciousness Studies, 5,* 389–404.

Camerer, C. (2006). Wanting, liking, and learning: Neuroscience and paternalism. *The University of Chicago Law Review, 73,* 87–110.

Camerer, C. (2007). Neuroeconomics: Using neuroscience to make economic predictions. *Economic Journal, 117,* C26–C42. doi:10.1111/j.1468-0297.2007.02033.x

Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature, 43,* 9–64. doi:10.1257/0022051053737843

Clark, A. (2011). *Supersizing the mind. Embodiment, action, and cognitive extension*. Oxford, UK: Oxford University Press.

Clark, L. (2010). Decision-making during gambling: An integration of cognitive and psychobiological approaches. *Philosophical Transactions of the Royal Society, 365,* 319–330.

Coleman, J. (1990). *Foundations of Social Theory.* Cambridge, UK: Belknap.

Dennett, D. C. (1991). *Consciousness explained.* Boston, MA: Little, Brown and Company.

Downs, J. S., Loewenstein, G., & Wisdom, J. (2009). Strategies for promoting healthier food choices. *American Economic Review, 99,* 159–164. doi:10.1257/aer.99.2.159

Elster, J., Ed. (1986). *The multiple self.* Cambridge, UK: Cambridge University Press.

Emmeche, C. (2002). The chicken and the orphan egg: On the function of meaning and the meaning of function. *Sign System Studies, 30,* 15–32.

Finlayson, G., Halford, J., King, N., & Blundell, J. (2010). *The regulation of food intake in humans.* Retrieved from http://www.endotext.org/obesity/obesity7.3/obesity7–3.html

Floridi, L. (2003). Information. In L. Floridi (Ed.), *The Blackwell guide to the philosophy of computing and information* (pp.40–61). Oxford: Blackwell.

Foxall, G. R. (2008). Reward, emotion and consumer choice: From neuroeconomics to neurophilosophy. *Journal of Consumer Behavior, 7,* 368–396. doi:10.1002/cb.258

Fudenberg, D., & Levine, D. K. (2006). A dual-self model of impulse control. *American Economic Review, 96,* 1449–1477. doi:10.1257/aer.96.5.1449

Glimcher, P. W. (2003). *Decisions, uncertainty, and the brain. The science of neuroeconomics*. Cambridge, MA: MIT Press.

Glimcher, P. W. (2009). Choice: Towards a standard back-pocket model. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 503–522). Amsterdam: Elsevier.

Glimcher, P. W. (2011). *Foundations of neuroeconomic analysis*. Oxford, UK: Oxford University Press.

Griffiths, M. D. (2008). The biopsychosocial and "complex" systems approach as a unified framework for addiction. *Behavioral and Brain Science 31,* 446f. doi:10.1017/S0140525X08004822

Gul, F., & Pesendorfer, W. (2008). The case for mindless economics. In A. Caplin & A. Schotter (Eds.), *The Foundations of Positive and Normative Economics* (pp. 3–41). Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780195328318.003.0001

Harrison, G. W. (2008). Neuroeconomics: A critical reconsideration. *Economics & Philosophy, 24,* 303–344. doi:10.1017/S0266267108002009

Herrmann-Pillath, C. (2010). Entropy, function, and evolution: Naturalizing Peircian Semiosis. *Entropy 12,* 197–242. doi:10.3390/e12020197

Herrmann-Pillath, C., & Salthe, S. N. (2011). Triadic conceptual structure of the maximum entropy approach to evolution. *BioSystems, 103,* 315–330. doi:10.1016/j.biosystems.2010.10.014

Hubert, M. (2010). Does neuroeconomics give new impetus to economic and consumer research? *Journal of Economic Psychology, 31,* 812–817. doi:10.1016/j.joep.2010.03.009

Hubert, M., & Kenning, P. (2008). A current overview of consumer neuroscience. *Journal of Consumer Behavior, 7,* 272–292. doi:10.1002/cb.251

Hughes, J. R. (1989). Environmental determinants of the reinforcing effects of nicotine in humans. *Journal of Substance Abuse, 1,* 319–29.

Hull, D. L., Langman, R. E., & Glenn, S. S. (2001). A general account of selection: Biology, immunology, and behavior. *Behavioral and Brain Sciences, 24,* 511–573.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics, 37,* 1555–1577. doi:10.1016/j.pragma.2004.06.008

Jablonka, E., & Lamb, M. J. (2006). *Evolution in four dimensions. Genetic, epigenetic, behavioral and*

symbolic variation in the history of life. Cambridge, UK: MIT Press.

Jamison, J., & Wegener, J. (2010). Multiple selves in intertemporal choice. *Journal of Economic Psychology, 31,* 832–839. doi:10.1016/j.joep.2010.03.004

Kahnemann, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives, 20,* 3–24. doi:10.1257/089533006776526030

Kahnemann, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of Experienced Utility. *Quarterly Journal of Economics, 112,* 375–405.

Kåhre, Jan (2002). *The Mathematical Theory of Information.* Boston: Kluwer.

Knutson, B., Delgado, M., & Phillips, P. E. M. (2009). Representation of subjective value in the striatum. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 389–406). Amsterdam: Elsevier.

Knutson, B., & Greer, S. (2008). Anticipatory affect: Neural correlates and consequences for choice. *Philosophical Transactions of the Royal Society B, 363,* 3771–3786. doi:10.1098/rstb.2008.0155

Knutson, B., & Wimmer, G. E. (2007). Reward: Neural circuitry for social valuation. In E. Harmon-Jones & P. Winkielman (Eds.), *Social neuroscience* (pp. 157–175). New York, NY: Guilford Press.

Krohs, U., & Kroes, P. (Eds.) (2009). *Functions in biological and artificial worlds*. Cambridge, MA: MIT Press.

Laibson, D. (2002). A cue-theory of consumption. *The Quarterly Journal of Economics, 116,* 81–119. doi:10.1162/003355301556356

Landreth, A., & Bickle, J. (2008). Neuroeconomics, neurophysiology and the common currency hypothesis. *Economics & Philosophy, 24,* 419–429. doi:10.1017/S0266267108002058

Lau, J., & Deutsch, M. (2008). Externalism about mental content. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2008 ed.). Retrieved from http://plato.stanford.edu/archives/fall2008/entries/content-externalism/

Lea, S. G., & Webley, P. (2006). Money as tool, money as drug: The biological psychology of a strong incentive. *Behavioral and the Brain Sciences, 29,* 161–209. doi:10.1017/S0140525X06009046

Le Folla, B., & Goldberg, S. R. (2005). Control of the reinforcing effects of nicotine by associated environmental stimuli in animals and humans. *Trends in Pharmacological Sciences, 26,* 287–293. doi:10.1016/j.tips.2005.04.005

Lende, D. H. (2008). Addiction: more than innate rationality, behavioral and brain science (commentary on Redish et al.) *Behavioral and Brain Sciences, 31,* 453–454.

Macdonald, G., & Papineau, D. (2006). Introduction: Prospects and problems for teleosemantics. In G. Macdonald, & D. Papineau (Eds.), *Teleosemantics. New philosophical essay* (pp. 1–22). Oxford: Oxford University Press.

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007). Time discounting for primary rewards. *The Journal of Neuroscience, 27,* 5796–5804. doi:10.1523/JNEUROSCI.4246-06.2007

Mesoudi, A., Whiten, A., & Laland, K. N. (2006). Towards a unified science of cultural evolution. *Behavioral and the Brain Sciences, 29,* 323–383. doi:10.1017/S0140525X06009083

Mick, D. G. (1986). Consumer research and semiotics: Exploring the morphology of signs, symbols, and significance. *Journal of Consumer Research, 13,* 196–213. doi:10.1086/209060

Mick, D. G., Burroughs, J. E., Hetzel, P., & Brannen, M. Y. (2004). Pursuing the meaning of meaning in the commercial world: An International review of marketing and consumer research founded in semiotics. *Semiotica 152,* 1–74.

Mijovic-Prelec, D., & Prelec, D. (2010). Self-deception and self-signalling: A model and experimental evidence. *Philosophical Transactions of the Royal Society, B, 365,* 227–240.

Millikan, R. (1989). Biosemantics. *Journal of Philosophy, 86,* 281–297. doi:10.2307/2027123

Morewedge, C. K., Huh, Y. E., & Vosgerau, J. (2010). Thought for food: Imagined consumption reduces actual consumption. *Science, 330,* 1530–1533. doi:10.1126/science.1195701

Neander, K. (2006). Content for cognitive science. In G. Macdonald, D. Papineau (Eds.), *Teleosemantics. New philosophical essays* (pp. 167–194). Oxford: Oxford University Press.

Neander, K. (2009). Teleological theories of mental content. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2009 ed.). Retrieved from http://plato.stanford.edu/archives/win2009/entries/content-teleological/

Oullier, O., & Kelso, J. A. S. (2009). Social Coordination, from the perspective of coordination dynamics. In R. A. Meyers (Ed.), *Encyclopedia of complexity and systems science* (pp. 8198–8212). Berlin: Springer.

Oullier, O., Kirman, A. P., & Kelso, J. A. S. (2008). The coordination dynamics of economic decision making: A multilevel approach to social neuroeconomics. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 16,* 557–571. doi:10.1109/TNSRE.2008.2009960

Peirce, C. S. (1992). *The Essential Peirce. Selected Philosophical Writings, Vol. 1 (1867–1893).* Indianapolis, IN: Indiana University Press.

Peirce, C. S. (1998). *The Essential Peirce. Selected Philosophical Writings, Vol. 2 (1893–1913).* Indianapolis, IN: Indiana University Press.

Platt, M., & Padoa-Schioppa, C. (2009). Neuronal representations of value. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 441–462). Amsterdam, The Netherlands: Elsevier.

Redish, A. D., Jensen, S., & Johnson, A. (2008). A unified framework for addiction: Vulnerabilities in the decision process. *Behavioral and Brain Sciences, 31,* 415–487. doi:10.1017/S0140525X0800472X

Reid, M. S., Ho, L. B., & Berger, S. P. (1996). Effects of environmental conditioning on the development of nicotine sensitization: Behavioral and neurochemical analysis, *Psychopharmacology, 213,* 301–310. doi:10.1007/BF02247381

Reimann, M., & Bechara, A. (2010). The somatic marker hypothesis as a neurological theory of decision-making: Review, conceptual comparisons, and future neuroeconomics research. *Journal of Economic Psychology, 31,* 767–776. doi:10.1016/j.joep.2010.03.002

Richerson, P. J., & Boyd, R. (2005). *Not by genes alone. How culture transformed human evolution.* Chicago, IL: University of Chicago Press.

Robinson, A., & Southgate, C. (2010). A general definition of interpretation and its application to origin of life research. *Biology and Philosophy, 25,* 163–181. doi:10.1007/s10539-009-9188-4

Robson, A. J., & Samuelson, L. (2010). The evolutionary foundation of preferences. In J. Benhabib, M. Jackson, & A. Bisin (Eds.), *Handbook of social economics* (pp. 221–310). Amsterdam: North-Holland.

Ross, D. (2005). *Economic theory and cognitive science: Microexplanations*. Cambridge, MA: MIT Press.

Ross, D. (2007). H. sapiens as ecologically special: What does language contribute? *Language Studies, 29,* 710–731.

Ross, D., Sharp, C., Vuchinich, R. E., & Spurrett, D. (2008). *Midbrain mutiny. The picoeconomics and neuroeconomics of disordered gambling.* Cambridge, MA: MIT Press.

Ross, D., Spurrett, D., Kincaid, H., & Stephens, G. L. (Eds.) (2007*). Distributed cognition and the will. Individual volition and social context.* Cambridge: MIT Press.

Saad, G. (2007). The evolutionary bases of consumption. New York, NY: Lawrence Erlbaum.

Salthe, S. N. (2009). The system of interpretance: Meaning as finality. *Biosemiotics, 1,* 285–294. doi:10.1007/s12304-008-9023-3

Schantz, R. (Ed.) (2004). *The externalist challenge*. Berlin, Germany: De Gruyter.

Schultz, W. (2009). Midbrain Dopamine Neurons: A Retina of the Reward System? In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 323–329). Amsterdam: Elsevier.

Simmons, W. K., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cerebral Cortex, 15,* 1602–1608. doi:10.1093/cercor/bhi038

Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure.* Cambridge, UK: Cambridge University Press.

Skyrms, B. (2010). *Signals. Evolution, learning, & information,* Oxford, UK: Oxford University Press.

Smith, V. L. (2003). Constructivist and ecological rationality in economics. *The American Economic Review, 93,* 465–508. doi:10.1257/000282803322156954

Sobel, J. (2005). Interdependent preferences and reciprocity. *Journal of Economic Literature XLIII,* 392–436. doi:10.1257/0022051054661530

Sobal, J., & Wansink, B. (2007). Kitchenscapes, tablescapes, platescapes, and foodscapes. Influences of microscale built environments on food intake. *Environment and Behavior, 39,* 124–142. doi:10.1177/0013916506295574

Sprevak, M. (2009). Extended cognition and functionalism. *Journal of Philosophy CVI,* 503–527.

Stanton, A. A., Welpe, I. (2010). Risk and ambiguity: Entrepreneurial research from the perspective of economics. In A. A. Stanton, M. Day, & I. M. Welpe (Eds.), *Neuroeconomics and the firm* (pp. 29–49). Cheltenham: Edward Elgar.

Steedman, I., & Krause, U. (1986). *Goethe's Faust, Arrow's Possibility Theorem and the Individual Decision-Taker,* In J. Elster (ed.), The multiple self (pp. 197–231), Cambridge: Cambridge University Press.

Sterelny, K. (2004). Externalism, epistemic artefacts and the extended mind. In Schantz, R. (Ed.), *The externalist challenge* (pp. 239–254), Berlin: de Gruyter.

Stone, T. L. (2007). *Peirce's theory of signs*. Cambridge, UK: Cambridge University Press.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge. Improving Decisions about Health, Wealth, and Happiness*. London, UK: Penguin.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science, 315,* 515–518. doi:10.1126/science.1134239

Tooby, J., & Cosmides, L. (2005). *Conceptual Foundations of Evolutionary Psychology.* In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 5–67). Hoboken, NJ: Wiley Publisher.

Trepel, C., Fox, C. R., & Poldrack, R. A. (2005). Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. *Cognitive*

Brain Research, 23, 34–50. doi:10.1016/j.cogbrainres.2005.01.016

Vehkavaara, T. (2002). Why and how to naturalize semiotic concepts for biosemiotics. *Sign System Studies, 30,* 293–313.

Wagner, D., Dal Cin, S., Sargent, J. D., Kelley, W. M., & Heatherton, T. F. (2011). Spontaneous action representation in smokers when watching movie characters smoke. *Journal of Neuroscience, 31,* 894–898.

Wansink, B., Just, D. R., & Payne, C. R. (2009). Mindless eating and healthy heuristics for the irrational. *American Economic Review, 99,* 165–169. doi:10.1257/aer.99.2.165

Wegner, D. M., & Sparrow, B. (2007). The puzzle of coaction. In D. Ross, D. Spurrett, H. Kincaid, & G. L. Stephens (Eds.), *Distributed cognition and the will. Individual volition and social context* (pp. 17–38). Cambridge, MA: MIT Press.

West, R. (2006). *Theory of Addiction*. Oxford, UK: Blackwell and Addiction Press.

Wilcox, N. T. (2008). Against simplicity and cognitive individualism. *Economics & Philosophy, 24,* 523–533. doi:10.1017/S0266267108002137

Wilson, R. A. (2004). *Boundaries of the Mind. The Individual in the Fragile Sciences: Cognition.* Cambridge, UK: Cambridge University Press.