Naturalizing Institutions:
Evolutionary Principles and Application on the Case of Money

by
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
Email c.herrmann-pillath@fs.de

Abstract

In recent extensions of the Darwinian paradigm into economics, the replicator-interactor duality looms large. I propose a strictly naturalistic approach to this duality in the context of the theory of institutions, which means that its use is seen as being always and necessarily dependent on identifying a physical realization. I introduce a general framework for the analysis of institutions, which synthesizes Searle's and Aoki's theories, especially with regard to the role of public representations (signs) in the coordination of actions, and the function of cognitive processes that underlie rule-following as a behavioural disposition. This allows conceiving institutions as causal circuits that connect the population-level dynamics of interactions with cognitive phenomena on the individual level. Those cognitive phenomena ultimately root in neuronal structures. So, I draw on a critical restatement of the concept of the 'meme' by Aunger to propose a new conceptualization of the replicator in the context of institutions: the replicator is a causal conjunction between (physical) signs and neuronal structures which undergirds the dispositions that generate rule-following actions. Signs, in turn, are outcomes of population-level interactions. I apply this framework on the case of money, analysing the emotions that go along with the use of money, and presenting a stylized account of the emergence of money in terms of the naturalized Searle-Aoki model. In this view, money is a neuronally anchored sign for emotions relating with social exchange and reciprocity. Money as a meme is physically realized in a replicator which is a causal conjunction of money artefacts and money emotions.

Keywords: Generalized Darwinism; institutions; replicator/interactor; Searle; Aoki; naturalism; memes; emotions; money

JEL classification: B52; D02; D87; E40; Z1

1. The naturalistic turn in the evolutionary approach to institutions

One of the major challenges in generalizing the theory of evolution is to include human culture and institutions into the picture (Mesoudi et al. 2006; Mesoudi 2011). In economics, this research agenda was launched by Thorstein Veblen (1899) for the first time, but was lost out of sight for most of the 20th century. Outside economics, the co-evolution of human biology and culture received considerable attention in anthropology and biology after sociobiology had attacked the foundations of the social sciences and humanities as independent research traditions. Today, diverse approaches to gene-culture evolution are at hand which avoid fully-fledged reductionism but also extend the evolutionary concepts into the realm of culture (e.g. Richerson and Boyd 2005; Jablonka and Lamb 2006); the concept of 'inclusive inheritance' is emerging as a unifying framework (Danchin et al. 2011). In economics, unified approaches are only back on the research agenda with the recent program of 'Generalized Darwinism' (Hodgson 2002; Aldrich et al. 2008; Hodgson and Knudsen 2010).

One fundamental conceptual problem in all these extensions is the question of ontology, in the specific sense of social ontology. Social ontology defines the major ontological difference between 'old' and 'new' institutionalisms in economics: In approaches of the former, seminally launched by Veblen, amongst others, institutions are treated as constituent units of social reality, whereas new institutionalisms mostly follow the standard assumption of methodological individualism in economics, which would only treat 'individuals' as 'real' units of larger social systems (for an overview, see Hodgson 1999). For evolutionary approaches and Darwinism in particular, this disjunction applies as well, in the context of the tensions between claims of genetic reductionism and the possible role of alternative approaches which would highlight the role of higher-level units in evolution. I reduce these complexities to one question: Can we construct an extension of evolutionary theory that treats institutions as units of evolution, alongside with genes as units of biological evolution? How can we elaborate on the general hypothesis on the 'ontological continuity' between different levels of evolution (Witt 2003)?

In this paper, I present an argument in favour of treating institutions as 'real' and as units of evolution on an ontological level which is independent from the genetic level. As such, the paper picks up a distinction which is currently seen as obsolete by the vast majority of researchers, namely the distinction between genes and 'memes', which was posited in one early universalization of Darwinism by Dawkins (1989) (for a collection of viewpoints, see

Aunger 2000). The problems in fixing the concept of the 'meme' are just special examples of the troubles with generalizations of another Dawkins concept that underlies the notion of meme, i.e. the 'replicator' (Hull and Wilkins 2005), and which has been put at the centre of the efforts of Knudsen and Hodgson (Knudsen 2002; Hodgson and Knudsen 2006). In particular, what is the physical realization of replicators in the institutional domain, beyond chemical mechanisms such as genetics?

I will put together different theoretical resources from different disciplines to offer my solution to this quandary. The first starts out from Aunger's (2002) theory of (neuro)memes. Aunger, too, posits that one of the intricate questions of the generalization of Darwinism is the distinction between replicators and interactors, i.e. the generic conceptual counterpart to the genotype / phenotype dualism. Replicators would define the information accumulation, transmission and retention function; interactors would define the functions of this information relative to selective environments. Although it is possible to make sense of this distinction in purely information-theoretic terms (as in Hodgson and Knudsen 2010, following related views such as Dennett 1995), this approach is methodologically problematic because it implicitly gives up the naturalistic ontology underlying Darwinism. In fact, the purely information-theoretic approach is a disguised Cartesian substance dualism of mind vs. matter which builds the universalization of the theory on the distinction between a material domain, where biology reigns, and an abstract domain of 'information', in which the generalization holds (for a pertinent discussion on the notion of information in biology, see Maynard Smith 2000 vs. Griffiths 2001; also compare Oyama 2001). Instead of this, I present an entirely naturalistic account of institutions (for a related view, see Sperber 2011). I define naturalism as the ontological hypothesis that the world is physically closed in causal terms, thus eschewing any sort of substance dualism, and that therefore 'existence' is defined in terms of physical causal powers (Papineau 2007, 2009). Naturalism does not preclude the possibility of emergence, i.e. ontological novelty (Bunge 1977/1979); that means, I propose a non-reductionist evolutionary account of institutions (in the sense of Hodgson 1999).

My naturalistic approach focuses on the causal circuitry between artefacts and neuronal structures as the physical realization of replicator functions. Then, one central question is how we can understand the causality between artefacts and neuronal states. My solution to this problem is to synthesize the categories of 'meaning' and 'function' in an evolutionary account of institutions. This synthesis starts out from the recent 'cognitive turn' in institutional economics (e.g. North 2005), which sees institutions as combinations of incentive mechanisms and cognitive schemes (mental models etc.). I present a detailed proposal on how to conceptualize replicators in the context of cognitive theories of institutions, taking Aoki's (2007, 2010, 2011) theory as a workhorse. Aoki has shown that for the emergence and sustainability of institutions a specific kind of causal circuitry between external artefacts (his 'public representations') and states of individuals (his 'beliefs') is essential. This causal circuitry mediates between individual-level and population level processes. I show that this view can be translated into purely naturalistic terms (compare Sperber 2011). As a side effect, my argument points toward a lacuna in recent debates in evolutionary economics, namely the integration of the brain sciences and recent progress in neuroeconomics. So, I propose an extension of the cognitive approach to institutions into a neurocognitive foundation, following recent theories about 'grounded cognition' and related approaches which emphasize the essential role of externalized actions in enabling cognitive formations (e.g. Barsalou 1999, Pecher and Zwaan 2005).

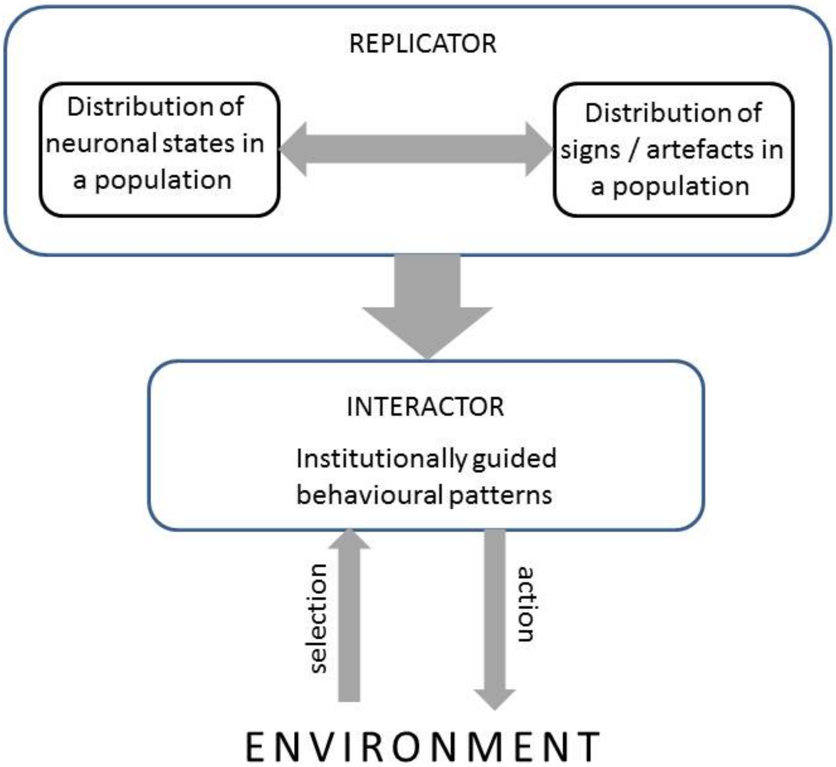
This argument basically follows the example of Hayek, who had put the analysis of the brain at the centre of his entire approach to institutions, starting out from his seminal 'Sensory Order' (Hayek 1952). As a result, I argue that the evolution of human institutions takes place at the interface of two levels of evolutionary processes, namely the evolution of states of the brain (Neural Darwinism, as launched by Edelman 1987) and the evolution of signs in the most general sense. These two processes connect with the process of genetic evolution via epigenetic mechanisms and the phylogenetic heritage of value functions that guide human learning in an institutional context.

Another central idea of the Hayekian approach to institutions that can be detailed analytically, too, in Aoki's framework is that institutions are media of distributed knowledge. Aoki analyses institutions in terms of a specific causal circuitry that relates interactions under institution with external sets of public representations, which have the essential function of 'information compression'. This idea matches exactly with the information-theoretic interpretation of the replicator / interactor duality. In Aoki's conceptualization, the signs have the function of information compression, and they generate certain dispositions that result into actions which reproduce certain behavioural regularities as well as the public representations. Thus, we get an empirical interpretation of the general replicator function in the context of a fully-fledged evolutionary approach to institutions. The replicator is a conjunction of signs and neuronal states, and the interactor is the resulting behaviour, however in terms of its aggregate, population level patterns, i.e. the 'states of play' in Aoki's sense. This analysis catches the important fact that both the interactor and the replicator must be population-level phenomena, such as in the classical distinction between the genotype and the phenotype

(Lewontin 2007). Even though the neuronal states are strictly individual, the signs are population level phenomena, and their functional relation depends on the sustainability of collective behavioural patterns in the population of agents. I call these patterns ‘institutionally guided behavioural patterns’ IGBP; thus, I highlight that behaviour is always individual behaviour in ontological terms, yet manifests collective properties because of the existence of institutions.

Thus, in summary, I present a Darwinian account of institutions that interprets institutions in terms of the replicator-interactor duality. The interactor is the pattern of sustainable behavioural regularities on the population level which manifest functions relative to a selective environment. The replicator is a stable causal conjunction of signs, which are generated on the population level, and neuronal states on the individual level. The replicator connects signs and individual behavioural dispositions, such that meaning of the signs is the function that it has in sustaining the population-level patterns. I summarize this basic structure in figure 1.

Figure 1: Replicator and interactor in the evolutionary approach to institutions



The paper proceeds as follows. In section two, I present a detailed account of the theory sketched so far. In section three, I apply this theory on the institution of money, putting together three different sources of insights: first, the empirical record of the role of emotions in the societal use of money; second, a specific proposal by Lea and Webley (2006) about the core emotion that undergirds the use of money, which I interpret in Darwinian terms, namely the human instinct of social exchange; and third, a conceptual model of the historical emergence of coins presented by Hutter (1994), which I analyse in terms of my generic neurocognitive model. Section four summarizes the argument in terms of stating a general replicator / interactor structure for institutional analysis and glimpses the larger research agenda of a naturalistic theory of institutions.

2. Institutions, distributed cognition and neuromemes: Outline of a naturalistic approach to institutions

In this section I develop the theoretical framework in more detail. I will relate different theoretical resources, and I present my own interpretations of these contributions. This is especially true for the pivotal theory, Searle's theory of institutions. I will introduce many of Searle's concepts, but impose a strictly naturalistic interpretation. Searle himself is arguing in naturalistic terms (e.g. Searle 2004), but also maintains what I call a 'mentalistic' approach, or, in other words, internalism with regard to mind. Especially, I focus on one notion that has been retreating in Searle's own work recently (e.g. 2010), which is the 'background'. The background is a set of enabling capacities of agents, which makes rule following possible. In his earlier (1995) work Searle argues that the background generates behavioural dispositions (which I distinguish sharply from actions or intentions). This idea I will relate with another theory of institutions in economics, which I find congenial to Searle's approach, namely Aoki's, especially in its most recent versions. For Aoki's approach, too, I develop a naturalistic account, and I will show that the linkage between the two theories rests on the notion of a functional causal circuitry, mediated via language (more generally, sign systems) and dispositions, such that institutions emerge as population-level regularities in individual behavioural patterns.

2.1 Searle: Institutional facts and functions

To begin with, Searle (1995: 129ff; compare Searle 2004) argues that institutionalized behaviour builds on behavioural dispositions, which are neurophysiologically anchored. ‘Following a rule’ does not require knowing the rule as such, so there is no need for a fully-fledged mental representation. It suffices to be able to process environmental cues which trigger neurophysiological reactions that produce the required behaviour. Thus, in this view institutions are not fully reflected in cognitive models, but in complex conjunctions of partial cognitive representations and neurophysiological dispositions. This viewpoint seems to be complementary to Aoki’s (2001) notion of the stabilization of institutions by summary representations of the underlying game structures. Summary representations are partial cognitive models which do not need to be shared in a population (contra Denzau and North 1994), but form part and parcel of the reproduction of the institution by means of coordinated behaviour, as long as pay-offs stabilize both the different summary representations and the behaviour. In elaborating on this model, Aoki (2007, 2010, 2011) introduces the notion of ‘substantive institutions’. This compares with the mentalism of many theories about institutions, especially in game theory. Mentalism approaches institutions as coordinated states of mind between individuals, especially in the sense of mutually confirming expectations, based on common knowledge. To the contrary, substantive institutions are external determinants of mental states, i.e. beliefs. I argue that this approach can be directly connected to Searle’s theory of institutions as facts, which is in turn based on a theory of functions. This theory of functions is central to understand the replicator-interactor relationship in the complete model that I propose: A replicator is a special kind of function relative to the interactor, and the interactor realizes functions relative to the environment.

Searle puts his theory in the broader context of a general theory of facts (Searle 1995: 120ff.). He distinguishes between observer independent and observer relative facts (in 2010, he changes the terminology from ‘observer’ to ‘mind’, which I do not follow here). A metal coin is a piece of metal, which is observer independent. But the function as money is entirely dependent on the observer, hence observer relative. Observer relativity ultimately roots in collective intentionality. This is a crucial step, if we further consider that Searle distinguishes between two kinds of the subjective / objective distinction, i.e. the ontological one, referring to facts about entities, and the epistemological one, referring to judgments about facts (see table 1). A fact can be ontologically subjective and epistemologically objective, such as in case of a technological artefact, which is observer relative, but the functioning of which follows physical laws. On the other hand, there can be ontologically objective facts which are epistemologically subjective, such as the so-called qualia, i.e. inner perceptions of feelings, which are physical states of the brain, but nevertheless cannot be directly accessed by outside observers. From these distinctions, it becomes clear that institutions are ontologically subjective but epistemologically objective. So, money is a part of an ontology which is observer relative, but its functionings can be analysed by objective epistemic tools, such as the quantity theory of money.

Table 1: Types of facts and examples

judgement / entity	Epistemically subjective	Epistemically objective
Ontologically subjective	Subjective fact	Institutional fact
Ontologically objective	psychoneural fact (qualia)	Biological fact

There are further important distinctions, especially between *agentive* and *non-agentive functions*, and, on the level of institutions, *regulatory* and *constitutive* ones. Agentive functions involve intentional agents not only in the ascription of the function, but also in its workings. That is, the function of my heart is non-agentive because its works independent from my intention. If I use money, this function depends on me and all other agents who agree to use money. However, in institutional analysis many functions are also non-agentive, if there are also collectively unintended consequences which might be only perceivable to the external observer. Regulatory institutions refer to a pre-existing field of activity, such as institutions governing the exploitation of fish resources; constitutive institutions create the very activity that is governed by the institution, as in the case of a financial market. Money can be regarded as a constitutive institution in the case of modern money, whereas the transition from pure commodity money to coins starts out from regulatory institutions. Further, agentive functions can result into states in which the process of collective intentionality actually retreats into the 'background', such as when coins as money are taken for granted. This implies a shift from agentive to non-agentive functions. Actually, we can state that in the evolution of institutions, the transformation from agentive to non-agentive functions is the essential process in the general phenomenon of institutional scaffolding of individual behaviour (North 2005).

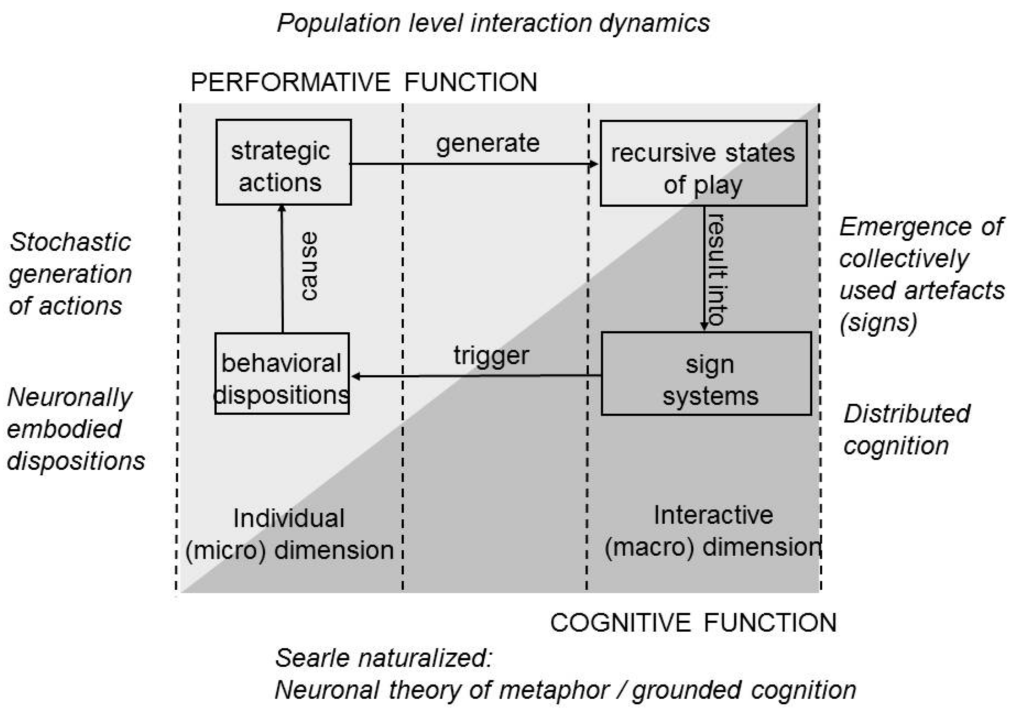
2.2. Aoki: Signs and dispositions in the emergence and reproduction of institutions

This analysis is essential to develop on a radicalization of Aoki's theory (for more detail, see Herrmann-Pillath 2012). This is because in Aoki's original argument, deployed in 2007 and 2011, there are still some traces of mentalism in two senses. One is that the public representations are seen in terms of a semantics of representation (survey in Lycan 1999), and the other is that they still generate 'beliefs' qua mental states. I propose to substitute this with a functional semantics which follows recent developments in teleosemantics (overview in MacDonald and Papineau 2006 and Neander 2009; compare Millikan 2005, 2009). This means to analyse the public representations exclusively in terms of their causal effects in the causal circuitry of institutions. To indicate this change of perspective, I use the term 'sign' instead of 'public representations', also following recent developments in game theory (e.g. Skyrms 2004, 2010). A sign does not 'represent' a state of play, but has a function, which consists of triggering certain responses by the agents that in turn support those states of play dynamically, which includes the production of those signs. This causal circuitry is the ontological feature that justifies treating institutions as facts in the naturalistic sense.

Then, the functions of institutions are partly independent from mental states of the agents, because they are partly non-agentive on the population level. Treating institutions as combinations of agentive with non-agentive functions grasps the Aoki concept of information compression, i.e. there can be no full assignment of functions by any purely internal mental states of individual observers because they are lacking the necessary knowledge of doing that. In other words, following an institution is normally based on a certain understanding of an institution, but also includes many unintended effects which are essential for the functioning of the institution, and which enable the generic function of information compression. The causal circuitry of the revised Aoki model describes an externalist approach to institutions, which allows arguing that via the institutions cognitive functions are externalized on population level processes.

I summarize the revised Aoki model in fig. 2. This keeps the original distinction between the individual(micro) and the aggregate population(macro) level, but changes the original orthogonal distinction between the behavioural and the cognitive level into the two notions of *distributed cognition* and *performativity*, which actually connect the individual level with the macro-level (indicated by the diagonal) (for more detail, see Herrmann-Pillath 2012). I also maintain the idea that there are strategic interactions in populations which can be analysed by different tools familiar from game theory and other approaches in economics and complex systems sciences (top of diagram). These interactions result into states of play which include the generation of signs or sign systems. Signs are artefacts which may be partly produced intentionally, but their functioning in the causal circuitry does not rely on this property (i.e. the function is non-agentive in essence). This is essential to understand the role of signs in processes of distributed cognition. Next, the signs produce causal effects that are mediated via neuronal structures of the agents. Thus, I merge the two notions of function and meaning in the sense that the meaning of the sign is its function relative to the neuronal states of the agents. The neuronal states create dispositions to act, again, without the essential requirement of consciously reflected choices. Dispositions cause actions within a certain range of random variation, which renders the entire model evolutionary in the Darwinian sense. The actions of different agents play together on the population level, producing certain aggregate results, possibly including the reproduction of the sign systems.

Figure 2: The Aoki model of institutions, naturalized



Thus, in case of sustainability we can say that the function of the signs is the maintenance of the behavioural regularities (which, interestingly, corresponds to the Wittgensteinian notion of meaning as context-bound action, contrary to representationalist approaches to meaning, see Lycan 1999). This is reflected in a twofold merger of the individual, micro-level phenomena, and the aggregate macro-level phenomena: First, cognitive functions are distributed across individual neuronal states and sign systems on the population level, and second, because of the causal relation between signs and dispositions, I argue that the actions become performative in the sense that rule-following behaviour emerges in which the individual actions anticipate the collective-level results within a certain range, thus reproducing them. In this view, collective intentionality in the sense of Searle's is supervening on the causal conjunction between signs and dispositions.

This naturalistic approach to institutions ultimately confronts population-level processes related to the emergence and diffusion of signs on the one hand with neuronal processes on the other hand. The bridge between the two levels is created via the cognitive functions of signs. This allows offering a naturalistic account of Searle's central theoretical concept in the analysis of institutions, the 'status function'.

2.3 Neurocognitive foundations of Searlian status functions

Searle's status function builds on language in a most general sense. In a status function, a certain entity is treated as another entity in functional terms; hence a metaphorical relation is created, depending on a particular context. This status function has the general form:

<*X* counts as *Y* in context *C*>

For example, I can treat a piece of metal as 'money', which is different from just using that piece in a simple barter process in which the traders may have some generalized use for it. A status function involves a fundamental conceptual shift to another category of meanings. Institutions presuppose collective intentionality in the sense that the status function must build on a collectively shared understanding in mutually connected actions (for an extensive account of this relationship, see Tuomela 1995, 2007, and Tuomela 2011 with reference to Searle). If in a community of language users a certain metal is used as money, single individuals cannot change this use just by taking an autonomous decision (along Wittgensteinian 1958 lines of his 'private language argument', see Candlish 2004).

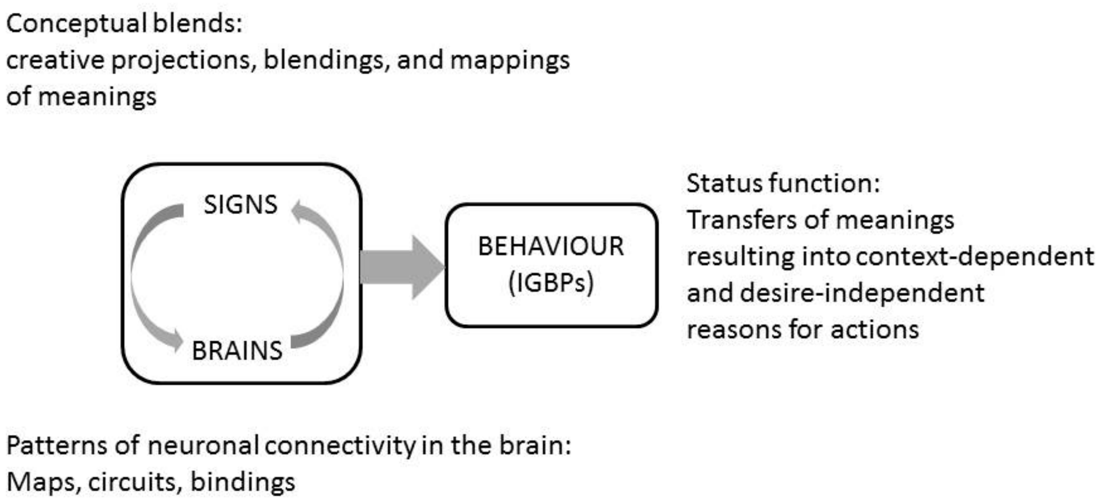
The status function is a metaphor in the most general sense: There is a transfer of meaning across previously disconnected domains, which is the very foundation of institutional creativity. Human individuals collectively create a new concept, money, for which a piece of metal, the coin, serves as a metaphor (this cognitive shift is overlooked in individualistic deconstructions of Searle's analysis of money, such as Smit et al. 2011). I will now argue that Searle's status function can be interpreted in terms of a causal conjunction between signs and neuronal states, thus completing the naturalistic revision of the Aoki model (for more detail, see Herrmann-Pillath 2010, 2010/11: 126ff.). Lakoff (2008) has developed a neuronal theory of metaphor which can help to clarify the linkage between the notion of the status function, which is a purely cognitive notion, and the notion of a neurophysiologically rooted behavioural disposition.

This theory identifies structural and processual characteristics of the brain which are supposed to correspond to the conceptual operations that underlie the creation and use of metaphors (Gentner and Bowdle 2008). As it is also assumed in generalized connectionist and network models of cognition (for a seminal approach, see Strauss and Quinn 1997), concepts are seen as corresponding to arrangements of neuronal groups; these groups are overlapping and organize into higher-level circuits with different kinds of patterns of connectivity (for a survey of related empirical brain research, see Coulson 2008). The basic mechanism driving this pattern formation is the differential activation and inhibition of neurons that form part of a group that is demarcated by these peculiar conjunctions of dual effects of neuronal activity. Recurrent activation and inhibition linkages result in neuronal bindings of different degrees of rigidity; these patterns are reflected in synchronized neuronal firings across populations of neurons. In particular, moving balances between inhibition and activation can catalyse the activation of circuits across different domains in the brain, thus resulting into mappings across larger structures. Then, dominant groupings can be activated by a single sensory input (for example, the fear of snakes is activated by the view of a snake-like movement); in so-called Gestalt circuits, the perception of parts of a phenomenon is completed by internal constructive processes in the brain. Such Gestalt mechanisms also work in the cross-domain mappings, such that the partial activation of one group simultaneously activates the entire other group in other parts of the brain. For the analysis of the cognitive processes that underlie status functions in institutional theory, two larger structural features of connections in the brain are crucial, mappings and bindings. Maps are projections across different brain areas, bindings establish a unity between more simple constituent units which can also relate with independent concepts.

The dynamics of this system is guided by a small number of principles, of which the best-fit principle is very important for the evolution of conceptual structures (for a related brain theory highlighting the underlying statistical principles, see Friston 2010). The best-fit principle means that the brain maximizes connectivity under the constraint of inhibitory relations between neuronal substructures. This is the force that underlies Gestalt dynamics and the creation of meanings from contextual cues. As a result of these different dynamics, the brain can build more complex linkages between simple mappings and bindings, which in turn can be the object of further mappings and bindings. At the same time, if only for stochastic reasons, those structures can be continuously reshuffled and recreated.

Based on this generic neuronal theory of concept formation and metaphor, it is possible to argue theoretically that the higher-level patterns of cognitive functions reflect these fundamental connectionist mechanisms. Therefore, I propose to relate the notion of status function to Fauconnier and Turner's (2002; 2008) theory of conceptual blends, resulting into a three-level neurocognitive structure of institutions which I overview in figure 3. On the basic level, we consider neuronal dynamics and structures of individual brains. This relates with higher-level processes of concept formation and conceptual blending, which are enabled by mechanisms of distributed cognition; that means, on the one hand blending is a process that is embodied in neuronal connectionist structures, but on the other hand this involves the manipulation of external signs, as I have outlined in the previous section. This causal conjunction of signs and neuronal processes enables the emergence of behaviours which manifest collective-level properties, i.e. institutionally guided behavioural patterns. On the conceptual level, these can be described by status functions that reflect the collective intentionality in using certain conceptual blends that commit individuals to shared patterns of actions; in the Searlian parlance, this means that the reasons for action become independent from individual desires.

Figure 3: Three-level neurocognitive structure of institutions



In a conceptual blend two concepts in different conceptual spaces are blended within a certain generic space, so that a new concept emerges, which may show also a new blend of pre-existing properties (see below, figure 5, for visualization). A blend is a higher-order neuronal mapping that builds on more elementary maps in the neuronal system, and crosses different parts of the brain (compare Coulson 2008). Because of the highly fluid nature of neuronal group selection, one can explain why the brain constantly creates novelty in the sense of new mappings between partial aspects of concepts that organize sensory and motor inputs. In a selectionist system, there can be no fixed reference, and only fuzzy meanings pop up from the dynamic course of neuronal variation, selection and retention (Edelman 2006: 98ff.). Thus, conceptual evolution might be imagined to proceed from some most basic concepts to the more complex ones, even though this is only a merely analytical convenience, because for a certain set of concepts at a particular point of time, all concepts relate synchronically, independent from their presumed diachronic order of emergence.

There are different kinds of blends, with different degrees of structural complexity. All blends can be described as conceptual networks with directed relations. For example, a simplex network is just a projection of an abstract reference frame on a particular input space. The most interesting case is the double-scope network, which is also most relevant for our analysis of the status function. In a double-scope network, two input spaces are projected into a generic space, such that the blend does not retain all properties of the input spaces, that is, it emerges as a novelty not only as a concept, but also regarding the specific combination of manifest properties. Clearly, this allows for the construction of new blends if that first blend is then projected into other generic spaces with other inputs and so forth.

Following Hutchins (2005), it is important to notice that conceptual blends are not simply based on neuronal connectivity; they need to rely on external complements. This view needs further elaboration in terms of the basic model of neurocognitive processes: In which sense can we say that concept formation and conceptual blending necessarily have to rely on a causal circuitry that connects neuronal processes with external signs?

2.4. Neural Darwinism and externalism

Lakoff's approach can be put into the general framework of Neural Darwinism that was developed by Edelman (1987) for the first time and has been extended into a full scale theory of mind and cognitive processing (Edelman 2006), presaged in Campbell's (1960) seminal thinking, and even Hayek's (1952) early work (for other related seminal approaches, see Calvin 1996, 1998; Hull et al. 2001). The structure of Edelman's theory is as follows (for summaries, see Edelman and Tonioni 1995 and Sacks 1995, or Friston 2010 on the relationship to other general brain theories). The basic idea is that the brain operates as a group selectionist system; groups of neurons compete against each other in mobilizing resources for neuronal growth and activity, such as neurotransmitters or energy in the most general sense.

The evolution of the neuronal system starts with developmental selection before birth, enabled by the highly disordered growth of neuronal connections under genetic constraints which creates the random variation necessary for selection to operate. Presumably, neuronal capacities for primary emotions emerge, which have phylogenetic roots and shape further brain development in terms of a set of basic evaluative mechanisms, related to survival and reproduction (especially primary emotions such as rage or lust, see Toronchuk and Ellis 2005). In Lakoff's theory, and corresponding to Searle's framework, these basic structures build the basis for the transformation of primary bodily, i.e. sensorimotor experience into more abstract conceptual schemes, especially linguistic representations (Gallese and Lakoff 2005; Gibbs and Matlock 2008; for a critical review, see Pinker 2007: 235ff.).

After birth, experiential selection takes place through which connections between neurons are strengthened according to differential sensory and motor inputs. Selection is guided by the set of genetically transmitted value systems, which define the fitness of neuronal units as reflecting certain causal mappings between events and states of the world and their effects on survival and reproduction; these function via the release of a number of neurotransmitters and other chemical substances in the brain, such as endorphines which relate with affects of pleasure or, especially within the context of learning, dopamine, which relates values, expectations and realized actions (for an overview, see Glimcher 2009). Building on these elementary structures, a further increase of complexity is achieved by means of re-entrant signalling and re-entrant mapping, which corresponds to Lakoff's similar notions of mappings and bindings. This refers to the increasing density of signal relations between neuronal groups which map different aspects of reality. Re-entrant signalling is different from feedback mechanisms in neuronal network models of error correction, as it primarily results in the active construction of the world by the brain (Friston and Stephan 2006). Via re-entrant signalling, neuronal groups end up in stable arrangements.

Now, in order to complete the general framework for analysing the neuroscience foundation for the phenomenon of the institutionalization of behaviour, what is missing is the role of communication across brains (compare Oullier, Kirman and Kelso 2008; Oullier and Kelso 2009). This question is directly relevant for relating the analysis so far with the role of signs in the causal circuitry of institutions. Edelman (1987: 320ff.) had already stated that without signal exchange across neuronal systems, no internal stabilization would be possible. This amounts to a fundamentally externalist position in the philosophy of mind, which sees mental facts as causal conjunctions of neuronal states and external causal processes (on externalism in general, see Schantz 2004; on causal theories of mind see Neander 2009 and Kim 2009).

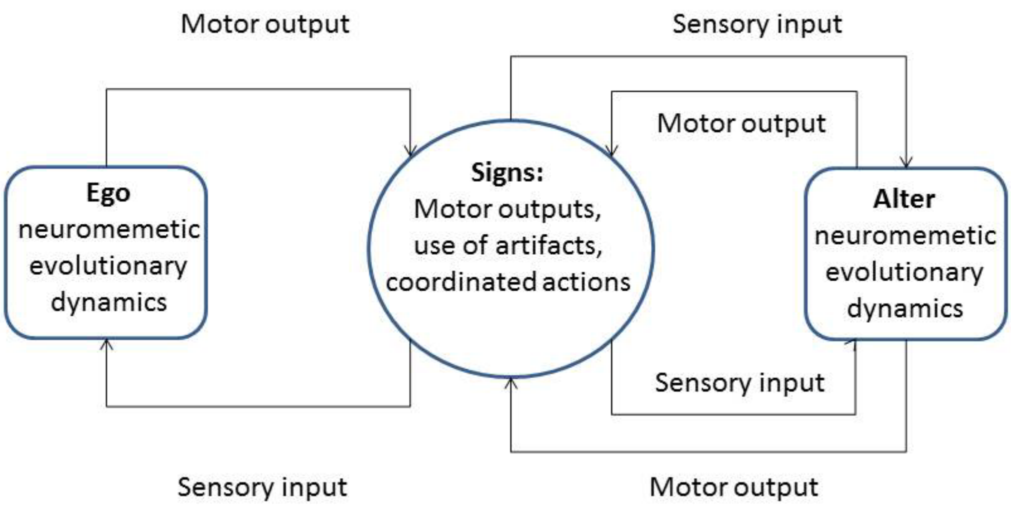
One conceptual approach to grasp this role of external factors is the theory of neuromemes that has been presented by Aunger (2002). In a nutshell, we can equate recurrently reproducing neuronal structures with 'neuromemes' which relate with external artefacts or signs, thus directly connecting with the causal circuitry analysed by Aoki. The neuromeme is a replicative neuronal structure within the human brain, which is defined according to certain structural effects in the on-going evolution of the neuronal architecture, both in the static and the dynamic sense. This proposal fits into Neural Darwinism framework. The specific mechanisms are still theoretical, but make empirical sense (see e.g. Fernando et al. 2008, Fernando et al. 2010).

By definition, neuromemes do not reside outside the individual brain, so that there is no way to presume that neuromemes are the same across different brains (which is a fundamental difference to genes, which migrate physically between bodies). This differs fundamentally from the original Dawkins conception of memes: The neuromeme is a unit of neuronal evolution, but not of cultural evolution. They do not have meanings in the sense of culture, but are defined according to neuronal functionings, as outlined in the Neural Darwinism approach. At the same time, neuronal evolution is a process that is basically independent from genetic evolution. That is, neuromemetics, following theories of neuronal selection and Neural Darwinism, posits that gene-culture coevolution is based on the simultaneous and interlocking runs of myriads of autonomous neuronal evolutionary sequences, with the neuromeme as essential part of the replicator. However, the extension to the notion of culture requires the introduction of another conceptual category. This is communication across brains and, more specifically, artefacts (Aunger 2002: 276ff.).

Cultural meaning supervenes on communication across brains, which operates via artefacts, in Aunger's parlance, or signs, as stated previously. The notion of a sign includes a broad range of physical phenomena, not only artefacts in the usual sense, but also, and foremostly, embodied signals such as the soundwaves of language or body movements. Signs are an essential part of a closed causal circuit which underlies the process of imitation in populations of brains. In recent theories about concept formation, the basic sensorimotor feedback loop between motor outputs and the resulting sensory inputs, continuously matched with phylogenetically rooted valuation mechanisms, is the elementary unit from which more complex structures emerge, mediated via specific neuronal structures, in particular mirror neurons (Arbib 2007; Fogassi 2011). On the one hand, via output inhibition this builds the basis for the internalization of functionings (Hurley 2008). On the other hand, outputs can simultaneously be inputs into different brains, thus enabling double track feedback loops between Ego's and Alter's Outputs.

I present a simple version of this causal circuitry in figure 4. The fundamental unit of the evolutionary emergence of meanings in brains are neuromemes; neuromemes, however, can only be replicators of neuronal structures, since there is no physical chain of causation connecting neuromemes across Ego's and Alter's brains. Therefore, meanings presuppose another ontological level, the level of signs. Signs are essential in individual concept formation and emerge from feedback loops between motor outputs and sensory inputs from own actions (for a seminal approach, see Barsalou 1999). However, this alone cannot result into stable causal conjunctions, as purely individual uses would succumb to entropic decay from arbitrary connectivity (which, again, leads back to Wittgenstein's 'private language argument'). It is essential that signs also feed into other brains: They trigger a feed-back loop on two levels. The first is that a sign generated by Ego is a sensory input to Alter that generates a motor output that reproduces the sign. This is accompanied by the proprioperceptive feedback mechanism (right inner circle). This two-level process underlies imitation and can be also analysed in the opposite direction, of course.

Figure 4: Coordination between brains via signs



I argue that this causal pattern defines the essence of the cultural replicator, or, in my parlance, of institutions as memes: Memes are causal conjunctions of signs and neuromemes. Building on Aunger's approach, this view avoids two problems of other interpretations of the replicator in the context of culture. The first problem is that we cannot reasonably identify a simple physical embodiment of replicators and the pertinent mechanisms if we only focus on the artefacts; after all, artefacts do not reproduce themselves. The second problem, however, is that if we rely on the single closest physical complement to culture, the neuronal structures, although we can envisage them to be physical embodiments of replication within brains, we cannot reasonably imagine a neuronal mechanism that physically connects these neuronal mechanisms across brains. My proposal solves the two problems, while being rooted in recent advances in neurosciences: Concept formation is based on sensorimotor feedback loops, thus necessarily involves externalized patterns of actions (Gallese and Lakoff 2005). This allows for signs to adopt an essential role in the reproduction of neuromemes, which in turn stabilizes patterns of sign use. From this follow that the entire structure sketched in figure 4 maps the replicator in cultural and institutional analysis. This viewpoint vindicates the more general externalist approach to mind, especially in the shape of theories on distributed cognition (Hutchins 1995; Clark and Chalmers 1998; Sterelny 2004; Clark 2011). The brain necessarily relies on a large and open range of external objects to achieve a stable equilibrium in what would be a chaotic fluctuation of neuronal firings otherwise.

2.5. Institutions and replicators

I can now close the circle of my argument centring on the revised and extended Aoki model of institutions. I posit firstly, that institutions can be equated with memes in the sense of stable conjunctions of signs and neuromemes that result into IGBPs. The emerging neuronal patterns provide the foundation for behavioural dispositions that can be activated by external cues, the signs, and which are amplified via endogenous Darwinian mechanisms of neuronal selection and pattern emergence. These internal mechanisms provide the foundation for metaphors, which is the basis for Searle's status function. Hence, the neuroscience framework can help to grasp three facts. One is that the neuronal system is creative; the second is that it involves external facts and hence is anchored in an objective ontology; and third, the bridge between internal and external processes is built via concepts that have neurophysiological correlates.

It is important to notice that the causal linkages shown in figure 4 also allow for a naturalistic conception of performativity that we have introduced in the context of the Aoki model (which differs from other uses of term in the social sciences, such as Butler 2010, but converges with recent theorizing about the material nature of routines in organizations, see D'Adderio 2008, 2011). Performativity is based on feedback loops between motor outputs and sensory inputs, and it becomes externally anchored via the diffusion of the sign correlates in a population of individuals that communicate via the exchange of artefacts. That is, performativity follows from the externalist approach to the mind/brain, combined with collective intentionality that emerges from the coordination of neuromemetic evolution and sign evolution.

So, we have now put together all necessary elements for a naturalistic theory of institutions. This can be finally summarized in terms of the replicator-interactor duality, as sketched in the first draft of figure 1. The central observation is that in the phenomenon of institutions, evolutionary processes on different levels interact, i.e. the population level and the neuronal level. This interaction can be conceived along the lines of the naturalized Aoki model, which I have extended to include Searle's theory of institutions. On the population level, we locate the interactor, which is the recurrent equilibria in states of interactions between individuals, resulting into IGBPs. These states include the generation of signs, which connect causally with neuronal structures in the individuals. Those structures are the physical embodiments of Searle's status functions, which we can analyse on different levels (figure 3), in turn, especially the conceptual level, which is accessible via cognitive sciences, linguistics and related disciplines. The neuronal dynamics creates behavioural dispositions to display certain actions that interact on the population level. The neuronally mediated causal link between signs and dispositions is the replicator; its causal structure is outlined in figure 4. Clearly, in the context of Generalized Darwinism, this model is a co-evolutionary model in which the environment (in this case, the signs) and somatic storages of information co-evolve (Jablonka and Lamb 2006). Especially, this approach matches with recent biological theorizing about niche construction in evolution, which posits the existence of 'environmentally mediated genotypic associations' (see Odling-Smee et al. 2003: 220ff, 245); my argument is formally homologous, but instead of referring to genes as one pole of that mediation I relate this idea to neuromemes.

3. The evolutionary approach to institutions at work: Darwinizing money

3.1. The pivotal role of emotions in the naturalistic approach to institutions

How can we translate the theory sketched in the previous section into a set of analytical tools that we can apply on real-world institutions? The central problem is to find a conceptual bridge that allows connecting lower-level neuronal mechanisms with higher level cognitive functions. I propose that this tool is the concept of 'emotions' (LeDoux 1997). In the basic explanatory structure of Neural Darwinism, valuations play an important role, as they connect the neuronal dynamics with stored information about past adaptive functionings. Damasio has proposed the theory of 'somatic markers' to account for this function (Damasio 1995, Bechara and Damasio 2005). I follow the assessment of Reimann and Bechara (2010) who argue that the concept of somatic markers is most promising in integrating different neuroscientifically informed theories about human behaviour, especially in the context of economics. Damasio's theory also advances the idea that those evaluative functions flow together in the phenomenon of 'emotions'.

This way of thinking closely relates to the expanding field of neuroeconomics. Therefore, in this section I build on recent research in neuroeconomics in order to present an exemplary evolutionary analysis of what is probably the most essential institution in the economy, money. This approach offers a response to Glimcher's (2011: 423f.) challenge. Glimcher ponders that money should be one of the priority research topics in neuroeconomics, because of its centrality in the incentive systems in recently evolved human economic systems. He speculates that there must be a specific neuronal structure that undergirds the human handling of money. I submit a proposal how that structure might look like. However, I do not posit a specific structure, but a mechanism that builds on the generic structures of the human brain. This is because in my externalist framework, the specificity of the institution of money does not lay in the specificity of internal structures of the brain, but in the specificity of the money artefacts, hence the causal circuitry between the money signs and the emotions qua neuronal structures. So, my discussion of the institution of money serves two purposes. One is to present an application of the theory, but the second is to further develop the theory in terms of advancing the notion of emotion as an empirically tractable expression of neuronal dynamics, as in Damasio's work.

Emotions are higher level coordinating neuronal mechanisms that relate with fundamental valuations which result from evolutionary processes on different levels, in turn. This is not the place where I can discuss the immensely complex literature on cognition and emotion (with special reference to the notion of rationality, see Pham 2004, 2007). So I just posit one particular position. This is to define emotions as framed affects, such that the notion of the frame, well recognized in economics now (Gintis 2006), can serve as the conceptual bridge between the notions of emotions and institutions, as we can say that frames are just another expression of the more general notion of the status function. In evolutionary psychology, emotions are seen as higher-order neuronal structures that coordinate sets of more elementary affectual circuits (Tooby and Cosmides 2005: 52ff.). As such, they are necessarily related with cognitive structures, in the sense that the coordination builds on pattern recognition in the environment. Thus, the emotional complex underlying fear of snakes is related with cognitive mechanisms of identifying snakes, including all possible transfers of meaning by metaphorical uses of the term etc. (for a nuanced discussion of this point, see Prinz 2009). These views closely relate with the current views about the modularization of the human brain and its cognitive mechanisms (e.g. Sperber and Hirschfeld 2004). Then, we can posit that emotions also undergird certain cognitive domains: A foremost example in recent work in evolutionary psychology is the identification of the social exchange domain, which relates with particular emotional and cognitive capacities of humans (Cosmides and Tooby 2005). I will heavily draw on these results.

This argument can be also inverted, in the sense that cognition presupposes affectual mechanisms which provide the ultimate standards of evaluation, which is, for instance, essential to select informational cues from the environment (Pham 2007: 161ff.). This is the claim put forward by Damasio's theory of somatic markers. From this follows, that cognitive approaches in institutional economics cannot work without taking emotions as a central category (for rare and early proposals, see Frank 1988 or Elster 1998). However, so far economics is missing a conceptual framework for doing this, as even in the majority of heterodox approaches a variation on the theme of rationality is maintained, mostly in the shape of bounded rationality, and emotions are rarely systematically explored (Phelps 2009).

I argue that the concept of 'emotion' is the missing conceptual link between the general evolutionary framework developed in the previous section and the analysis of specific institutions such as money, which I present in this section. That is, I posit that the fundamental replicator in an institution such as money is a causal conjunction of a sign and an emotion, the latter understood as a higher-level neuromeme in the sense of Aunger's. Therefore, I begin with analysing the emotions that go along with the societal use of money.

3.2. Money emotions

The best place to begin an exploration of money emotions is where the pinnacle of the economic notion of rationality seems to have been materialized, i.e. the modern financial markets, which have been the object of the conceptual struggle between the rational theory of finance and behavioural finance in recent years.

Looking at the activities involving the 'money professions', recent anthropological and sociological research has shown that financial trading is a far cry from being a purely rational-calculative concern (Zaloom 2003; 2004). Trading financial assets involves very strong feelings and requires a special emotional discipline, which, however, does not simply mean to supress emotions, which are absolutely necessary to raise the alertness and aggressiveness indispensable to successful trading. Making money often shows a resemblance to making sex, highly exciting, but also highly dangerous, in the sense of financial ruin or social and physical havoc (sexual diseases, unwanted pregnancy etc.) (Seabright 2004: 76). This is reflected in the language and the habits of traders' communities (Hassoun 2005). After all, the financial business is also highly gender-biased with a very pronounced male dominance (Klaes et al. 2007). Recently, those affectual underpinnings have been related to different testosterone levels in both male and female traders affecting their relative professional success and other aspects of the neuronal and hormone system (Maestripieri et al. 2009; Wargo et al. 2010).

Thus, 'rationality' in financial trading results from special emotional disciplines and techniques of self-management that both contain and exploit underlying affectual drives (Preda 2008: 918). This can be also seen in the larger context of historical sociology, which reveals how the investor as a particular kind of personality emerged in the 19th century (Preda 2005). In an even broader perspective, the emergence of modern capitalism was accompanied by strong reactions in religion and folk beliefs, often resulting in a demonization of money. Against this background, Max Weber's account of the rise of capitalism acquires a new meaning, as he had argued that Calvinism inspired its believers with a particular emotional stance towards money. As an expression of 'inner worldly ascetism', Calvinists were able to pair the acquisitive drive with abstention from lavishly spending it, thus launching the machine of accumulation in early capitalism.

These sketchy observations clearly underline what is evident from our everyday experience: Money causes strong emotions, and using and spending money has an emotional basis. Indeed, one of the important results of neuroeconomics is that money comes close to being a direct reinforcer. That means, money activates the same dopaminergic circuits in the human brain (more exactly, the mesolimbic system) as other items causing pleasure, such as beautiful faces (Camerer et al. 2005: 35, Phelps 2009: 240). This simple fact is exploited in the work of psychologists, as money can be directly used as a generalized reward without further modification (Knutson and Wimmer 2007: 159f.). Money triggers general reward mechanisms in the human brain, possibly even involving a so-called 'common currency' (Landreth and Bickle 2008).

The autonomous role of money as a reinforcer is related with important anomalies in behavioural economics and finance. For example, people loath the loss of cash dividends and do not net them out with capital gains, so cash seems to carry an additional value (see surveys of behavioural finance such as Van der Saar 2004 or Subrahmanyam 2007). This can be explained by complex conceptual constructions, such as the interaction between loss aversion and hyperbolic discount curves. A present cash loss is weighted relatively strong against a future capital gain, even though the two might be equivalent for a 'rational' decision maker. People organize their perceptions in different mental accounts for income and wealth, such that current income shows a disproportionally strong impact on consumption behaviour, as compared to the predictions of rational choice theory (surveyed in Akerlof 2007). Thus, people seem to need a special approach to manage money in the narrow sense, that is, cash, which is deeply grounded in social norms and expectations. Lack of control in spending money is often seen as a lack of self-control. The special meaning of cash in those systems of behavioural regulation can be also gleaned from the fact that credit cards seem to loosen those constraints, presumably because they have different effects in the context of hyperbolic discount curves, combined with loss aversion (Laibson 1997).

These observations are also related to the equity premium puzzle, which has been explained by moving reference points with regard to dividend payments, as opposed to capital gains. If reference points move, different degrees of loss aversion are implied, thus explaining the additional risk premium necessary to make the trades equal. If this explanation is valid, however, this implies that money activates the neuronal mechanisms underlying frame-dependent loss aversion (Knutson et al. 2008), whereas the more complex accounting constructs of capital gains and losses do not. The same monetary values trigger different neuronal responses, depending on their representation.

The clearest proof, also confirmed by brain research recently, for the independent reward triggered by money is money illusion. The existence of money illusion has been confirmed by psychological research beyond any doubt (Shafir et al. 1997). For example, people normally report higher satisfaction with deals involving higher nominal quantities of money to deals with lower ones, even if, in an inflationary setting, the real values are the same. In brain imaging studies, researchers could show that the reward circuits in the brain react much stronger with the higher nominal, yet identical real values; and they could even identify proportional levels of activities in the pertinent brain areas (Weber et al. 2009).

To summarize, recent research in behavioural economics and neuroeconomics has shown that money can be seen as a direct reinforcer in the same way as other goods. Obtaining and keeping money satisfies a want that is independent from the derived need for money to obtain other goods. This implies that our standard conception of money may be misled by the assumption that money is a general purpose exchange medium. Although this use of money is part and parcel of the modern institution of money, it may not be at the core of the social fact.

In other words, the transactional use of money might be a derived function, which is, however, prevalent in modern economies. So-called behavioural anomalies of money use can be explained by the fact that these two functions of money interact in social practice.

3.3. Money, signal selection and the human instinct to social exchange

The analysis of money emotions presented so far does not yet identify a specific emotion that could be triggered by the use of money. Psychologists Lea and Webley (2006) have recently proposed a 'drug and tool' theory of money, which presents a proposal of identification. The upshot of their argument builds on a large number of empirical observations of the kind that I have sketched in the previous section, that is, in their wording, money is a strong and universal incentive. Clearly, money is used as a tool in many contexts, that is as a mere transaction device, but at the same time it manifests very strong emotional reactions in other contexts, and can trigger strong drives of acquisitive and hoarding behaviour. To grasp this phenomenon conceptually, the authors introduce the notion of a 'perceptual drug' which differs from a mere psychoactive drug such as nicotine. This specific argument about the addictive roles of money is problematic, as it does not match with established theories about addiction, which do not make the distinction between psychoactive and perceptual drugs (see e.g. the authoritative overview of West 2006). In some theories, the psychoactive drugs are only a special case of the more general model of addiction which emphasizes internal malfunctioning of brain mechanisms, especially related with learning (e.g. Redish et al. 2008). One approach (further elaborated in Herrmann-Pillath 2012b) emphasizes the fundamental duality of wanting and liking in the human neuronal system and resulting psychological setup, that I referred to already in the previous section: If money activates the same dopaminergic circuits as other positively valued things, this refers to the so-called 'wanting' system, so there is a difference to the 'liking' system which refers to the actual consummation (Trepel et al. 2005; Camerer 2006; Brocas and Carrillo 2008; Berridge 2009). This distinction confirms the distinction between 'experience utility' and 'decision utility' that has been proposed in the psychological literature on economic choice (Kahnemann et al. 1997). The wanting system underlies the processes that guide anticipatory planning and expectations. Thus, there is a duality of rewards operating in the activities of planning and acting to acquire a certain benefit, and the rewards actually resulting from consuming the pertinent acts. In the neuroeconomic literature, this intermediation is reflected in the autonomous role of the dopamine circuits in guiding action as distinguished from consummation of actions, which implies that the dopaminergic circuits themselves are involved in creating the phenomenon of addiction. This simple connection emerges from the fact that dopamine signals code deviations between actions and results, such that positive results trigger further action (reward prediction errors, see Schultz 2009). Thus, the dopamine circuits can establish self-reinforcing feedback loops. Addiction would result from failures of the associated learning process, such that the wanting system runs astray and drives behaviour autonomously. In the context of money, the most appropriate illustration is gambling (Ross et al. 2008, Clark 2010). In gambling, the individual seeks the rewards of money, but succumbs to the mechanism which increases dopamine levels in her brain because of the on-going perception of failures which are close to the target (hence indicate improvements, in the sense of positive reward prediction errors), and which are even interpreted as indicators of the individual skills of the gambler (illusion of control).

Now, we can put these observations into the context of a general Darwinian approach. The phenomenon of malfunctioning that is apparent in addictive behaviour can be related to the Darwinian theory of signal selection as stated by Zahavi and Zahavi (1997). The dualism of wanting and liking underlies the human capacity to plan and has phylogenetic roots in the evolution of the mammalian brain (overview in McCabe 2008: 354ff.). This, in turn, functionally requires representation, that is, builds on language, and more general on symbols that intervene in the causal process of eliciting goal-oriented behaviour (for more on that in a general evolutionary argument, see Dennett 1995 or Millikan 1989, 2005, 2009). Therefore, different from purely instinctual or reflexive reactions, human behaviour is systematically built on the distinction between the sign and the object. Thus, an apple is both on object and its sign. The signs play the crucial role in anticipatory reward mechanisms that underlie the wanting system. With this insight, we can establish a direct conceptual linkage with our previous analysis of the role of signs in institutions.

The theory of signal selection implies that for arbitrary signs, the so-called handicap principle may apply, depending on the selective context. The handicap principle posits that the coordination of behaviour via signals may require the investment into costly signals which produces an adaptive disadvantage in the sense of natural selection (as opposed e.g. to sexual selection) (Dawkins 1989: 309ff.; Grafen 1990). Yet, precisely these costs make the signal functional, because otherwise it would be open to manipulation and cheating. That is, handicaps are truthful signals and because of that, are adaptive in terms of the universal currency of reproductive success. I posit that the handicap principle underlies also the malfunctioning of the human decision system, elaborating on a suggestion of Ascoli and McCabe (2006) in their comment on the Lea and Webley paper. Ascoli and McCabe ponder whether the argument may hold for all scarce goods. An excellent example is eating (Berridge

2009). In times of scarcity of food, certainly prevalent throughout most of human phylogenetic past, the signs of food become exceptionally important for behavioural choices. This implies that the signs will also play a crucial role in behavioural coordination. Accordingly, food use is also governed by signal selection, ending up in the many examples of ritualized and very expensive and elaborate food customs. If that is the case, however, the sign of food can also trigger malfunctioning. As such, it underlies the many dysfunctions of eating. People who devour sweets without limits do not actually consume the sweets, but the signs of sweets, in this interpretation.

So, extending on the Lea and Webley argument, the question is which fundamental human drive and need might underlie the use of money. I propose a slight, yet essential modification of their central idea: which is, that the money drug piggybacks on an instinct to trade that evolved out of the universal mechanism of reciprocal altruism. More generally, we can point to the emotional patterns underlying social exchange that have been identified by evolutionary psychology (Cosmides and Tooby 2005, Ermer et al. 2006). Evolutionary psychology argues that the human species manifests a peculiar emotional structure that enables humans to maintain complex networks of social exchange based on reciprocity. Though reciprocity as such is a universal biological phenomenon (Trivers 1985: Chapter 3; Noë et al. 2001), the human species excels in terms of the generalization and the scope of exchange patterns (Seabright 2004). In these relations, both competitive and cooperative relations occur, often simultaneously, as modelled in game theoretic approaches towards egoistic cooperation (for example, in hunting large game, epitomized in Rousseau's stag hunt, see Skyrms 2004). As a result, modern evolutionary approaches to the development of the human brain posit the 'social brain' hypothesis (e.g. Dunbar and Shultz 2007; Frith 2007). Following up to earlier versions of Machiavellian intelligence (Byrne 1995), this hypothesis states that the evolutionary more recent and innovative neuronal structures in the human brain are geared towards the organization and manipulation of social exchange. This evolutionary argument can be further supported by the observation that phylogenetically closer animals can also show distinctly human deviations from rational choice models, if they also live in highly interactive social groups (Santos and Chen 2009).

From that perspective, money emerged as a sign that triggers emotional responses related to social exchange in general. These are affects that relate with calculating mutual benefits across time, with detecting cheaters, or with perceiving mutual relations of indebtedness. Money mobilizes these emotional patterns, without actually satisfying them, which can only be done with executing the underlying exchanges (i.e. the 'liking' system). Yet, money triggers the same reward mechanisms (the ‘wanting’ system), which, accidentally, also seem to be activated in the entirely different setting of PD dilemmata (Knutson and Wimmer 2007: 166; Fehr 2009). This seems to go back to the fact that the perception of cues to cooperation is tantamount to the perception of gains, i.e. rewards. Indeed, PD dilemmata also manifest social exchange relations, as it is evident from considering repeated games, which is the reasonable assumption for primordial human groups and trading communities. Therefore, the historical record (considered below) that indicates the primary role of hierarchically embedded reciprocity in the emergence of money seems to match with the observation that evolutionary more ancient patterns of social exchange might not have been related with market-kind behaviour, but with exchange of contrived goods (Ofek 2001: Chapter 9). The complexity of exchange in these cases results precisely from the intermingling of exchange relations with cooperative behaviour, as in maintaining and sharing fire, or in hunting and sharing large game.

Can we relate these insights to the historical analysis of the emergence of money? Clearly, early money was a phenomenon related with social exchange. Against the background of my evolutionary model of institutions, could we do some work in cognitive archaeology in order to show how the emotional underpinnings of the institution of money relate with the cognitive mechanisms that resulted with the diffusion of the artefact of money?

3.4. The emergence of money as the creation of a new metaphor

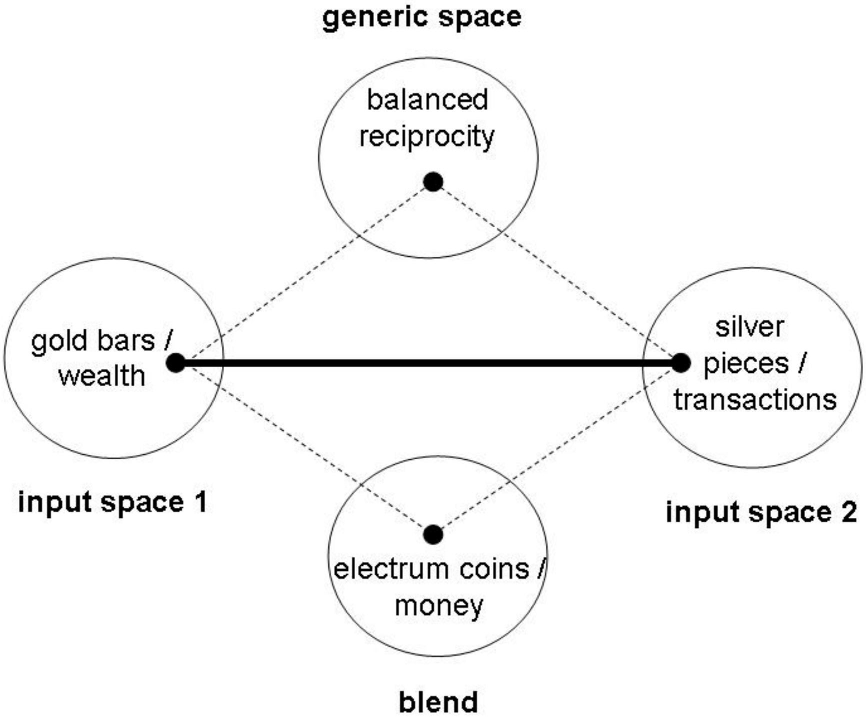
I will now use the notion of conceptual blending to reconstruct the historical emergence of the institution of money. Subsequently, I use money as a most generic term, that is, I do not differentiate between different kinds of money. However, I draw a relatively neat line between the mere fact of a particular good to be used as medium of exchange (pure commodity money) and money as an institution, which I relate to the emergence of the first coins. This transition is the central concern of this section. So, if I use the term ‘money’ subsequently, it is always in the second sense.

In this context it is important to notice that there is a clear contradiction between established theories of the emergence of money and the historical facts. Purely theoretical accounts emphasize the role of money in enabling transactions, as in the triangular exchange paradigm. The classic, almost unsurpassed until today, is Menger’s (1892) evolutionary account (for modern receptions, see e.g. Schotter 1982). In these accounts a certain item evolves as (commodity) money for the pure function of serving as a transaction device. This also implies that basically, there is no conceptual transformation, in the sense of a status function, but only a growing functional salience of properties such as resaleability, storability and dividability. In Searle's parlance, the institution of money would emerge as a purely regulatory institution, referring to commodity money.

This view clearly fits into the established economic theory of money, but contradicts the historical evidence. In fact, money emerged in the context of relations of power and authority, regulating hierarchical exchange relations and the production of public goods, and only the further evolution was intermingled with the more haphazard use of valuable items in barter (for a survey of the evidence, see Chavas and Bromley 2008). The historical data suggest that the transactional function of money is in fact a derived function. Hutter (1994) presents an intriguing account of the historical facts that we can directly translate into the conceptual blend framework and into the status function structure, which precisely models that derivative relation between pristine money and derived functions. In Searle's framework, those status functions would come close to a constitutive institution, in which money appears to be a new kind of thing, which actually superseded the parallel use of commodities for exchange.

When money coins emerged for the first time in the Eastern Mediterranean (at least as far as Western civilization is concerned), this was an effect of cross-cultural merger of meanings between Assyrian culture and the Ionian peasant communities. In Assyria, gold served as an indicator of status and as a medium of wealth accumulation in a steeply stratified society. In the Greek communities, silver was used for ritual purposes and occasionally for exchange, which was mostly mediated via a number of items with less value in barter, thus corresponding to the Menger view. The first genuine coins originating from Lydia, however, were made from electrum, an alloy of silver and gold. Thus, they could be interpreted differently in the two societies, enabling cross-cultural exchange of signs and goods (compare Hutchins 2005 on the role of material artefacts in enabling cognitive blends). Further, in order to test the quality of coins, people applied punches resulting into punch marks, firstly unintended. Once the coins circulated, people discovered the possible use of the punch marks as indicators of origin. The question of origin was crucial in lowering quality uncertainty with regard to the actual metal composition. From this moment onwards, the custom of coining emerged, with the incipient use of the punch marks as signals. Hutter (1994) speaks of an oscillation between the notions of 'signed metal' and 'metal sign'. Soon, the new coins were reintegrated into the political and the religious realm when local regents adopted the institution of minting. Thus, the first coins appeared displaying the images of rulers and holy symbols. With this transition, we can say that the institution of money has emerged.

Figure 5: Conceptual blending in the emergence of coins



Hutter's account easily can be translated into the framework of conceptual blends (figure 5). Simultaneously, we can apply the status function notion. As we see in a standard Fauconnier and Turner notation, the central point is that two different artefacts with different uses in different societies were merged into a common frame. This corresponds to a so-called 'double-scope integration network', in which two concepts are only partially merged, as in the case of 'time as space' (see Fauconnier and Turner 2008, Turner 2008). In a similar fashion, in the cross-cultural semantic ambiguity an alloy could be treated both as being close to a gold bar or a silver piece, thus also making those two artefacts commensurable, though only partially. In terms of the status function, in a double-scope integration network we can say that the status function in fact works in both directions, with 'gold' being treated as 'silver' and vice versa, in the context of cross-cultural exchange, and being related to the 'brute fact' of the physical entity of electrum. In other words, the physical fact of an alloy enabled the creation of the status function. Once the status function was established, the institution of money, reified in the emergence of the first coins circulating in inter-regional trade, came into existence. The new coins were clearly different from the mere use of commodities in exchange, hence constituting new things as artefacts.

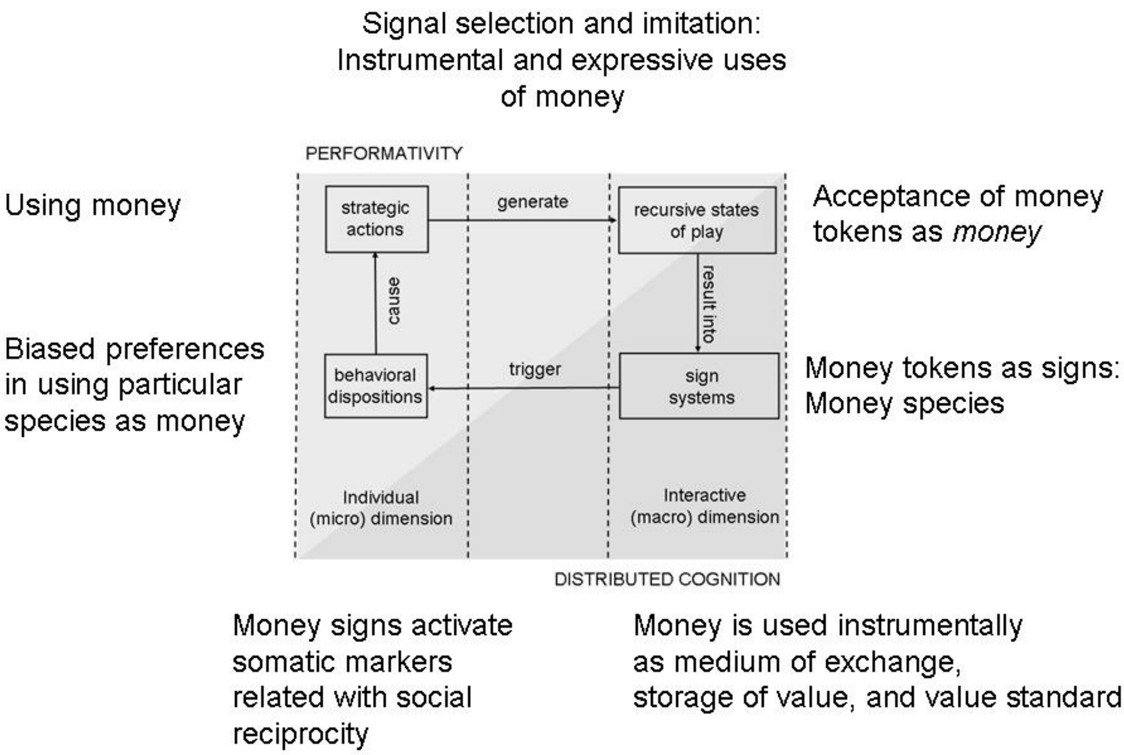
The question is what is the generic space that corresponds to the 'brute fact' of the alloy? And how can we relate this cognitive process to emotive structures? In figure 5, I propose to use the notion of 'balanced reciprocity', which is both an abstract notion underlying ritualized power relations and of early exchange across longer time horizons (a classic on this is Pryor 1977: Chapters 4 and 7; compare Burkert 1996: Chapter 6). That means, money and balanced reciprocity are deeply interconnected, which seems to be an acceptable intuition, which, however, turns into a hypothesis based on much broader evidence if we connect the status function model with the analysis of money emotions in the previous sections. Then, we can venture the hypothesis that the first coins evolved into metaphors for the notion of generalized exchange and reciprocity, which connects with the underlying emotions identified by evolutionary psychology.

So, if money is a sign of that special kind, it is also open to the functioning of signal selection, especially with regard to the handicap principle. So it is straightforward to explain why the emergence of money immediately went along with the appearance of extreme forms of accumulation and wealth display, as in Greek tyranny. In the original conceptual blend, money used for transactions and money used for storing and accumulating wealth were merged into one blend. This allows for a handicap mechanism to emerge, in the sense that the capability to hoard and display wealth is a marker of the capacity to spend limitless. The waste of wealth in useless displays is precisely the signal that communicates the capacity to enter a limitless number of transactions, thus claiming the role of a hub with exceptionally high prestige in ever-growing networks of exchange. This account matches the historical data in the sense that the purely technical uses of money emerged as a side effect of the primordial uses. Later, the runaway evolution of money also supported its rapid diffusion as a tool for transactions. This blend was epitomized in the emergence of the coin as an artefact for trade, but at the same time as a symbol for power and authority.

3.5. Causal circuitry in the emergence of money

I can now reconstruct the emergence of money in terms of the Aoki model, which specifies the underlying causal circuitry (fig. 6). As in the empirical case of Lydian commerce, resulting from cross-cultural interactions between traders, certain physical entities, the coins as money tokens, were increasingly used as money, that is, they were recognized socially as a new kind of artefact. This involved the transition from the tokens to the signs, hence led to the emergence of money species, i.e. certain classes of coins with similar characteristics. These social practices result into the realization of the distributed cognition function of the institution, i.e. the diffusion of money enables the innovation of new uses of the signs qua artefacts, such as in the different dimensions of money functions, understood in the traditional way. The use of particular money species is rooted in the status function which is in turn anchored in money emotions, centring on money as a metaphor for social exchange. This anchor creates a disposition to use money on the individual level, finding expression in particularly strong valuations of money. The individual use of money plays together with other uses on the population level, which establishes the performativity of the institution of money.

Figure 6: The emergence of money



We can now state that money is a meme, in the specific sense of the causal conjunction of certain signs, the money tokens circulating in a population, and the emotion related with social exchange, i.e. a particular neuronal structure. Interestingly, this structure can be traced back to the functioning of genetic evolution as well, as we can safely assume that there is direct phylogenetic line leading from the human ancestors to the modern human species, which included the particular somatic structures underlying the emotions of social exchange and reciprocity. However, it is not possible to reduce the cultural phenomenon of money to the genetic level, but only the underlying human penchant for social exchange. With the emergence of money, a fact on a different ontological level had emerged, which depends on the workings of the Searlian status functions.

As money is related with neuronal patterns underlying social exchange, mutually reinforcing causal feedback circuits emerged that further stabilized those human capacities, and open up the way for new expressions. Thus, with the diffusion of money its possible use as a transaction device was further strengthened, which in turn changed the context of social exchange towards the settings of more anonymous market-type relations. It is not the evolution of markets that required the emergence of money, but to the opposite: The evolutionary emergence of an artefact with the properties of early money made the further growth of markets possible, which is exactly the shift towards derivative functions of money. We can add that this transition is the crucial step towards performing a monetary exchange economy, with money becoming an observer relative fact, hence adding to the social ontology. Money as an externalization of social exchange emotions allowed for the emergence of new cognitive powers, such as the intersubjectively accepted calculation of values. This further enhanced its autonomous status in social reality, triggering behavioural innovations which were not possible without its existence (the classic on this, unsurpassed until today, is Simmel 1900; cf. Walsh and Lynch 2008; money compares with writing in this regard, on the latter see Menary 2007).

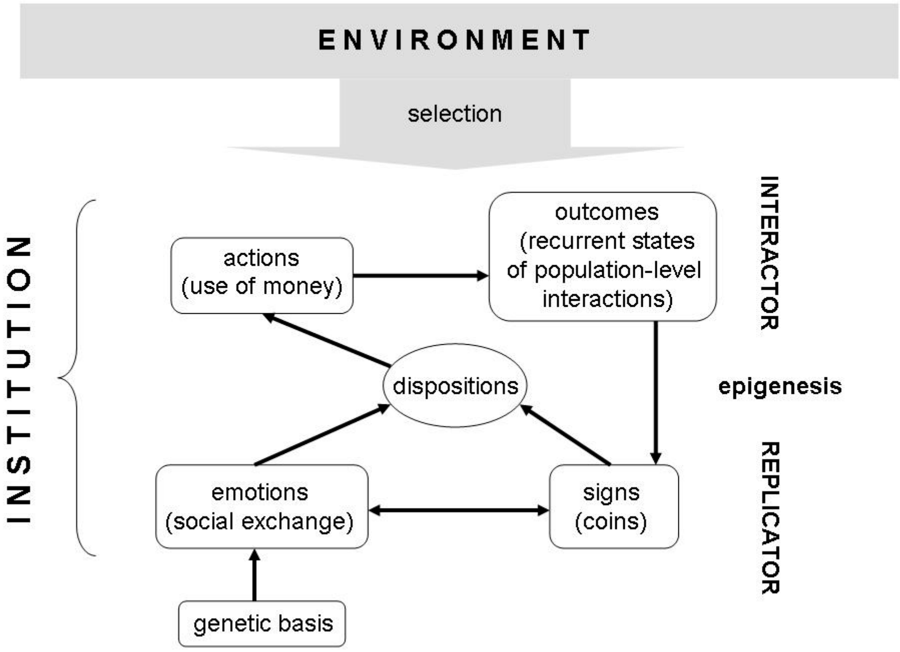
The naturalistic approach enables us to make sense of the empirical observations about the strong emotional components in handling money even in most advanced human societies. This does not imply, however, that those components always prevail. The artefact of money evolved historically in increasingly complex ways, thus strengthening functional and technological interdependencies. Therefore, actual money use today is governed by a mix of determinants, including also the 'rational' use of money without significant impact of primordial emotions. However, this means that those rational uses of money are not an outcome of the rationality of the human agent, but reflect the workings of external artefacts, such as institutions governing money markets, technologies governing money calculations in accounting, or new money artefacts such as electronic money. That is, the proposed perspective on money is also an externalist account of economic rationality in the modern uses of money. Evolving artefacts may trigger "irrational uses" (such as in the credit card case, Laibson 1997) or "rational uses" (such as with the evolution of modern accounting systems, see Hatherly at al. 2008). So, the Darwinian theory of money also fits into the conceptual schemes that are emerging in economic sociology as an extension of social studies on science, especially in the context of social studies on finance (Preda 2008). Here, agency on financial markets is increasingly seen as resulting from complex networks of interaction between individual behaviour and embedding technologies (for pertinent collections of papers, see Callon et al. 2007; Pinch and Swedberg 2008). This is just a special expression of the general neuromeme-artefact conjunction that was identified in this paper. So I expect that the framework can be extended to include also the more sophisticated institutions in modern finance.

4. Conclusion and outlook

I summarize the results of this paper in figure 7, focusing on the replicator-interactor duality, and summing up this paper in terms of adding analytical detail to figure 1. The first important insight is that we have to refer the term 'institution' to the entire causal circuitry, which corresponds to Aoki's model. This directly implies that we can think of the selection of institutions, relative to a particular environment. This selection works via the state of nature impact on the outcomes of the interactions within a population that follows an institution. Treating the institution as the entire causal circuitry corresponds to the type/token distinction in the relation between the individual and species: The individual corresponds to the individual action, and the species corresponds to the population level patterns. Insofar as following an institution is a performative action, the individual action reflects the population-level patterns, in the same way as the individual reflects the species features, within a range of variation, which, as we have seen for the case of institutions, result from the stochastic relationship between dispositions and actions.

Since selection operates via the outcomes of the interactions, we can identify the realized patterns of behaviour on the population level as the interactor in the generalized Darwinian sense. This catches the different uses of the term 'institution' in the literature, which vacillate between the rules of the game and the realized equilibria (e.g. Aoki 2001: 24ff; Dixit 2004: 5ff.). In my approach, both are true to a certain extent, but more exactly, the game equilibria are the interactor (which corresponds to the notion in biology that the interactor is actually equal to the selectively relevant behavioural patterns of a species).

Figure 7: The Darwinian approach to institutions



The interactor relates with the replicator via the process of epigenesis (compare Jablonka and Lamb 2006). I relate this process directly to the disposition that generates the individual actions. It is essential to see that this disposition is created by the causal conjunction of the signs and the emotions in the replicator. Therefore, the notion of epigenesis captures gene-culture co-evolution in the specific sense that the emotions are genetically anchored, but the gene expression is always mediated culturally, i.e. via the signs that evolve entirely independently from the genetic level. This extends the notion of epigenesis from somatic mechanisms to extra-somatic mechanisms. By implication, it is not possible to reduce observed actions directly to the genetic basis. This offers a solution to the nature-nurture quandary, as all gene expressions are seen as culturally contextualized, such that culture is an ontologically autonomous channel of the transmission of biological information (which matches with approaches such as Richerson and Boyd 2005 or Odling-Smee's et al. (2003) notion of 'cultural niche construction').

Thus, institutions build on a new kind of replicator, which, as I have discussed extensively, consists of a causal conjunction of signs and neuronal structures. The intricate property of this replicator is that the signs are strictly population level phenomena, whereas the neuronal structures are strictly individual phenomena. Important properties of the institutional replicator therefore directly relate with the material stability of the signs. Signs change, too, because of the population level processes. But it is important to notice that those changes only become evolutionarily relevant if they result into altered dispositions which in turn generate actions that change the population level equilibria. It is perfectly possible that a new sign generates similar actions and hence reproduces a similar institution.

In this paper, this model of institutions has been applied on a core institution of the economy, money. My analysis shows that it is necessary to strip institutions of their technical complexities in modern societies in order to make their evolutionary foundations explicit. The reason lies in the reflexivity of status functions, which create new social facts continuously. As we have seen in the case of money, this means that money today is a very complex phenomenon, with different uses and expressions in different contexts. Money emotions may not count much for a central banker. In my analysis I have not scrutinized the process of growing complexity in the evolution of institutions, which is often driven by the reflective powers of human reason and imagination. However, this is just another expression of the ontological autonomy of institutions. If we want to understand the fundamentals, we have to go back to the prototypical institutions, as I have shown for the case of money. On the other hand, complexity also means that those institutions also may exert an indirect impact on institutional innovations that connect with them. So, financial markets may bear the imprint of the primordial institution of money (which is evident from the many effects of money emotions on their performance). I think that a similar analysis is possible for other primordial institutions, as I have argued earlier (Herrmann-Pillath 1994) for the case of property. In standard modelling approaches to animal conflicts (the Hawk-Dove game), the so-called Strategy of Bourgeois is evolutionarily stable, which builds on a particular emotional structure, that is, to invest more resources into the defence of a territory if one is incumbent than if one is intruder. In human culture, this emotional structure has been causally conjoined with the emergence of numerous signs that signal the relative positions of incumbents and intruders. This would be the archetypical situation of the emergence of private property, akin to my analysis of money in this paper (especially if combined with cognitive analysis such as Schlicht 1998). I think that similar evolutionary approaches can be specified for other elementary human institutions, such as the institutions of cooperation (following Skyrms 2004 and others).

With relation to the more general problem of the extension of Darwinism into the domain of human social systems, this paper has shown that we can draw on more recent abstract reformulations of Darwinism, especially the replicator-interactor duality, in order to achieve a conceptual unification. However, I emphasize that this is only possible if we adopt a radically naturalistic ontology which follows the rule that all entities that are posited must have a physical realization. We may not yet have discovered such realizations, similar to Darwin's position, when genes were still unknown. But we always need to present a reasonable speculation. This paper offers a speculation about the physical nature of the replicator in institutional evolution: The causal conjunction of neuronal structures and signs. I think that this hypothesis can be also extended into the analysis of human material culture, that is, consumption. This raises the additional question how far consumption activities are institutionalized (see Boldyrev and Herrmann-Pillath 2013). However, the basic point seems promising, namely to analyse the evolution of human consumption patterns in terms of replicators which are ontologically defined as causal conjunctions of emotions and signs. There is already a growing literature on the Darwinian analysis of consumption (e.g. Saad 2007) which makes the accomplishment of this task straightforward. In the end, I expect that we will be able to identify a set of universal human emotions that connect with a cultural cosmos of signs that we can further systematize into different domains of institutions and economic activities. Whereas the set of emotions is mainly in the state of evolutionary stasis, the evolution of signs drives the continuous emergence of behavioural novelties, which also change the context of the expressions of human biology.

References

Adams, Fred, Aizawa, Ken, "Causal Theories of Mental Content", The Stanford Encyclopedia of Philosophy (Spring 2010 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2010/entries/content-causal/>.

Akerlof, George A. (2007): The Missing Motivation in Macroeconomics, American Economic Review, 97(1): 5-36.

Aldrich, Howard et al. (2008): In Defence of Generalized Darwinism, Journal of Evolutionary Economics 18(5): 577-598.

Aoki, Masahiko (2001): Toward a Comparative Institutional Analysis, Stanford: Stanford University Press.

Aoki, Masahiko (2010): Corporations in Evolving Diversity. Cognition, Governance, and Institutions, Oxford: Oxford University Press.

Aoki, Masahiko (2007): Endogenizing Institutions and Institutional Change, Journal of Institutional Economics 3(1): 1-31.

Aoki, Masahiko (2011): Institutions as Cognitive Media between Strategic Interactions and Individual Beliefs, Journal of Economic Behavior and Organization, vol. 79: 20-34.

Ascoli, Giorgio A. / McCabe, Kevin A. (2006): Scarcity Begets Addiction, Behavioral and the Brain Sciences 29: 178.

Aunger, Robert (2002): The Electric Meme. A New Theory of How We Think. New York et al.: Free Press.

Aunger, Robert, ed. (2000): Darwinizing Culture: The Status of Memetics as a Science, Oxford: Oxford University Press.

Ayres, Clarence E. (1944): The Theory of Economic Progress, Chapel Hill: University of North Carolina Press.
Barsalou, Lawrence W. (1999): Perceptual Symbol Systems, Behavioral and the Brain Sciences 22: 577-660.
Bechara, Antoine / Damasio, Antonio R. (2005): The Somatic Marker Hypothesis: A Neural Theory of Economic Decision, Games and Economic Behavior 52: 336-372.
Bentley, R. Alexander / Shennan, Stephen J. (2003): Cultural Transmission and Stochastic Network Growth, American Antiquity 68(3): 459-485.
Berridge, Kent C. (2009): ‚Liking' and ‚Wanting' Food Rewards: Brain Substrates and Roles in Eating Disorders, Physiology & Behavior 97: 537-550.
Bhaskar, Roy (1989): The Possibility of Naturalism. A Philosophical Critique of the Contemporary Human Sciences, New York et al.: Harvester Wheatsheaf.
Blackmore, Susan (2000): The memes' eye view, in: Aunger (2000): 25-42.
Boldyrev, Ivan / Herrmann-Pillath, Carsten (2013): Hegel's “Objective Spirit”, Extended Mind, and the Institutional Nature of Economic Action, Mind & Society DOI: 10.1007/s11299-012-0111-3.
Brocas, Isabelle / Carrillo, Juan D. (2008): The Brain as a Hierarchical Organization, American Economic Review 98(4): 1312-1347.
Bunge, Mario (1977): Treatise on Basic Philosophy, Volume 3. Ontology I: The Furniture of the World. Dordrecht et al.: Reidel.
Bunge, Mario (1979): Treatise on Basic Philosophy, Volume 4. Ontology II: A World of Systems. Dordrecht et al.: Reidel.
Burkert, Walter (1996): Creation of the Sacred. Tracks of Biology in Early Religions, Cambridge and London: Harvard University Press.
Butler, Judith (2010): Performative Agency, Journal of Cultural Economy 3(2): 147-161.
Byrne, Richard W. (1995): The Ape Legacy: The Evolution of Machiavellian Intelligence and Anticipatory Interactive Planning, in: Goody (1995): 37-52.
Callon, Michel / Millo, Yuval / Muniesa, Fabian, eds. (2007): Market Devices. Malden et al.: Blackwell.
Calvin, William H. (1996): The Cerebral Code. Thinking a Thought in the Mosaic of Mind, Cambridge/London: MIT Press.
Calvin, William H. (1998): Competing for Consciousness. A Darwinian Mechanism as an Appropriate Level of Explanation, Journal of Consciousness Studies 5(4): 389-404.
Camerer, Colin (2006): Wanting, Liking, and Learning: Neuroscience and Paternalism, The University of Chicago Law Review 73: 87-110.
Camerer, Colin / Loewenstein, George / Prelec Drazen (2005): Neuroeconomics: How Neuroscience Can Inform Economics, in: Journal of Economic Literature Vol. XLIII: 9-64.
Campbell, Donald T. (1987/1960): Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes, in: Radnitzky / Bartley (1987): 91-114.
Candlish, Stewart (2004): Private Language, The Stanford Encyclopedia of Philosophy (Spring 2004 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2004/entries/private-language/>.
Chavas, Jean-Paul / Bromley, Daniel W. (2008): On the Origins and Evolving Role of Money, Journal of Institutional and Theoretical Economics 164(4): 624-651.
Clark, Andy (2011): Supersizing the Mind. Embodiment, Action, and Cognitive Extension, Oxford: Oxford University Press.
Clark, Andy / Chalmers, David J. (1998): The Extended Mind, Analysis 58: 10-23.
Clark, Luke (2010): Decision-making During Gambling: An Integration of Cognitive and Psychobiological Approaches, Philosophical Transactions of the Royal Society B 2010 365: 319-330.
Cordes, Christian (2006): Darwinism in Economics: Fraom Analogy to Continuity, Journal of Evolutionary Economics 16(5): 529-542.

Cosmides, Leda / Tooby, James (2005): Neurocognitive Adaptations Designed for Social Exchange, in: Buss (2005): 584-628.
Coulson, Seana (2008): Metaphor Comprehension and the Brain, in: Gibbs (2008): 177-194.
D'Adderio, Luciana (2008): The Performativity of Routines: Theorising the Influence of Artefacts and Distributed Agencies to Routine Dynamics, Research Policy 37: 769-789.
D'Adderio, Luciana (2011): Artifacts at the Centre of Routines: Performing the Material Turn in Routines Theory, Journal of Institutional Economics 7(2): 197-230.
Damasio, Antonio R. (1995): Descartes' Error. Emotion, Reason and the Human Brain. New York: G.B. Putnam's Son.
Danchin, Étienne / Charmantier, Anne / Chanpagne, Frances A. / Mesoudi, Alex / Pujol, Benoit / Blanchet, Simon (2011): Beyond DANN: Integrating Inclusive Inheritance Into an Extended Theory of Evolution, Nature Reviews Genetics 12: 475-486.
Dawkins, Richard (1989): The Selfish Gene. New edition. Oxford/New York: Oxford University Press.
Dennett, Daniel C. (1995): Darwin's Dangerous Idea. Evolution and the Meanings of Life, New York: Simon & Schuster.
Denzau, Albert T./North, Douglass C. (1994): Shared Mental Models: Ideologies and Institutions, in: Kyklos 47: 3-32
Dixit, Avinash K. (2004): Lawlessness and Economics. Alternative Modes of Governance, Princeton and Oxford: Princeton University Press.
Dunbar, Robin / Shultz, Susanne (2007): Understanding Primate Brain Evolution, Philosophical Transactions of the Royal Society B, 362: 649-658.
Edelman, Gerald M. (1987): Neural Darwinism. The Theory of Neuronal Group Selection. New York: Basic Books.
Edelman, Gerald M. (2006): Second Nature. Brain Science and Human Knowledge. New Haven and London: Yale University Press.
Ermer, Elsa / Guerin, Scott A. / Cosmides, Leda / Tooby, John / Miller, Michael B. (2006): Theory of Mind Broad and Narrow: Reasoning About Social Exchange Engages ToM Areas, Precautionary Reasoning Does Not, Social Neuroscience 1(3-4): 196-219.
Fauconnier, Gilles & Turner, Mark (2002): The Way We Think. Conceptual Blending and the Mind's Hidden Complexities. New York: Basic Books.
Fauconnier, Gilles / Turner, Mark (2008): Rethinking Metaphor, in: Gibbs (2008): 53-66.
Fehr, Ernst (2009): Social Preferences and the Brain, in: Glimcher et al. (2009): 215-232.
Fernando, Chrisantha / Goldstein, Richard / Szathmáry, Eörs (2010): The Neuronal Replicator Hypothesis, Neural Computation 22: 2809-2857.
Fernando, Chrisantha / Karishma, K.K. / Szathmáry, Eörs (2008): Copying and Evolution of Neuronal Topology, PloS ONE 3(11), doi: 10.1371/journal.pone.003775
Fogassi, Leonardo (2011): The mirror neuron system: How cognitive functions emerge from motor organization, Journal of Economic Behavior & Organization 77(1): 66-75
Frank, Robert H. (1988): Passions Within Reason. The Strategic Role of Emotions, New York/London: Norton.
Friston, Karl (2010): The Free-energy Principle: A Unified Brain Theory, Nature Reviews Neuroscience doi: 10.1038/nrn2787
Friston, Karl J. / Stephan, Klaas E. (2006): Free-energy and the Brain, Synthese 159: 417-458.
Frith, Chris (2007): The Social Brain? Philosophical Transactions of the Royal Society B, 362: 671-678.
Gallese, Vittorio / Lakoff, George (2005): The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge, Cognitive Neuropsychology 22(3 & 4): 455–479
Gentner, Dedre / Bowdle, Brian (2008): Metaphor as Structure-Mapping, in: Gibbs (2008): 109-128.

Gibbs, Raymond W., ed. (2008): The Cambridge Handbook of Metaphor and Thought, New York et al.: Cambridge University Press.
Gibbs, Raymond W., Jr. / Matlock, Teenie (2008): Metaphor, Imagination, and Simulation. Psycholinguistic Evidence, in: Gibs (2008): 161-176.
Glimcher, Paul W. (2009): Choice: Towards a Standard Back-pocket Model, in: Glimcher et al. (2009): 503-522.
Glimcher, Paul W. (2011): Foundations of Neuroeconomic Analysis, Oxford and New York: Oxford University Press.
Glimcher, Paul W. / Camerer, Colin F. / Fehr, Ernst / Poldrack, Russel A., eds. (2009): Neuroeconomics: Decision Making and the Brain, Amsterdam et al.: Elsevier.
Goody, Esther, ed. (1995): Social Intelligence and Interaction. Expressions and Implications of the Social Bias in Human Intelligence. Cambridge: Cambridge University Press.
Grafen, Alan (1990): Biological Signals as Handicaps, Journal of Theoretical Biology 144:517-546.
Griffiths, Paul E. (2001): Genetic Information: A Metaphor In Search of a Theory, Philosophy of Science 68: 394–412
Hackett, Edward J. / Amsterdamska, Olga / Lynch, Michael / Wajcman, Judy, eds. (2008): The Handbook of Science and Technology Studies, Cambridge and London: MIT Press.
Harrison, Glenn W. (2008): Neuroeconomics: A Critical Reconsideration, Economcs & Philosophy 24(3): 303-344.
Hassoun, Jean-Pierre (2005): Emotions on the Trading Floor: Social and Symbolic Expressions, in: Knorr Cetina and Preda (2005): 102-120.
Hatherly, David / Leung, David / MacKenzie, Donald (2008): The Finitist Accountant, in: Pinch / Swedberg (2008): 131-160.
Hayek, Friedrich August von (1952): The Sensory Order. An Inquiry into the Foundations of Theoretical Psychology. Chicago: University of Chicago Press
Herrmann-Pillath, Carsten (1994): Evolutionary Rationality, "Homo Economicus" and the Foundations of Social Order, in: Journal of Social and Evolutionary Systems, Vol. 17/1, 41-70.
Herrmann-Pillath, Carsten (2010): A Neurolinguistic Approach to Performativity in Economics, Journal of Economic Methodology 17(3): 241-260.
Herrmann-Pillath, Carsten (2010/11): The Economics of Identity and Creativity. A Cultural Science Approach, St. Lucia: University of Queensland Press and New Brunswick and London: Transactions.
Herrmann-Pillath, Carsten (2012a): Institutions, Distributed Cognition and Agency: Rule-following as Performative Action, forthcoming: Journal of Economic Methodology.
Herrmann-Pillath, Carsten (2012b): Towards an Externalist Neuroeconomics: Dual Selves, Signs, and Choice, Journal of Neuroscience, Psychology and Economics 5(1): 38-61.
Hodgson, Geoffrey / Knudsen, Thorbjørn (2010): Darwin’s Conjecture. The Search for General Principles of Social and Economic Evolution, Chicago: University of Chicago Press.
Hodgson, Geoffrey M. (1999): Evolution and Institutions. On Evolutionary Economics and the Evolution of Economics, Cheltenham: Elgar.
Hodgson, Geoffrey M. (2002): Darwinism in economics: from analogy to ontology, Journal of Evolutionary Economics 12(3): 259-282.
Hodgson, Geoffrey M. (2002): Darwinism in economics: from analogy to ontology, Journal of Evolutionary Economics 12(3): 259-282.
Hodgson, Geoffrey M. (2004): The Evolution of Institutional Economics: Agency, Structure and Darwinism in American Institutionalism, London: Routledge.
Hodgson, Geoffrey M. / Knudsen, Thorbjørn (2006): The Nature and Units of Social Selection, Journal of Evolutionary Economics 16(5): 477-490.
Hopfensitz, Astrid / Reuben, Ernesto (2009): The Importance of Emotions for the Effectiveness of Social Punishment, Economic Journal 119(540): 1534-1560.

Hull, David and John S. Wilkins, "Replication", The Stanford Encyclopedia of Philosophy (Fall 2005 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2005/entries/replication/>.
Hull, David L. / Langman, Rodney E. / Glenn, Sigrid S. (2001): A General Account of Selection: Biology, Immunology and Behavior, Behavioral and Brain Sciences 24(2): 511-573.
Hurley, Susan (2008): The Shared Circuits Model: How Control, Mirroring and Simulation Can Enable Imitation, Deliberation, and Mindreading, in: Behavioral and the Brain Sciences 31(1): 1-21.
Hutchins, Edwin (1995): Cognition in the Wild, Cambridge and London: MIT Press.
Hutchins, Edwin (2005): Material Anchors for Conceptual Blends, Journal of Pragmatics 37: 1555-1577.
Hutter, Michael (1994): Communication in Economic Evolution: The Case of Money Changing the Primary Distinction: From Observer to Observation, in: England, Richard W., ed., Evolutionary Concepts in Contemporary Economics, Ann Arbor: University of Michigan Press, 111-138.
Jablonka, Eva / Lamb, Marion J. (2006): Evolution in Four Dimensions. Genetic, Epigenetic, Behavioral and Symbolic Variation in the History of Life, Cambridge and London: MIT Press.
Kahnemann, Daniel et al. (1997): Back to Bentham? Explorations of Experienced Utility, Quarterly Journal of Economics 112: 375-405.
Kahnemann, Daniel et al. (1997): Back to Bentham? Explorations of Experienced Utility, Quarterly Journal of Economics 112: 375-405.
Kim, Jaegwon (2009): Mental Causation, in: McLaughlin (2009): 29-53.
Klaes, Matthias / Lightfoot, Geoff / Lilley, Simon (2007): Market Masculinities and Electronic Trading, SCEME Working Paper No. 14. www.sceme.rog.uk
Knorr Cetina, Karin / Preda, Alex, eds. (2005): The Sociology of Financial Markets, Oxford et al.: Oxford University Press.
Knudsen, Thorbjørn (2002): Economic Selection Theory, Journal of Evolutionary Economics 12(4): 443-470.
Knutson, Brian / Wimmer, G. Elliott (2007). Reward: Neural Circuitry for Social Valuation. In Harmon-Jones, E., Winkielman, P. (eds.) Social Neuroscience, New York: Guilford Press, 157-175
Knutson, Brian et al. (2008): Neural Antecedents of the Endowment Effect, Neuron 58: 814-822.
Laibson, David (1997): Golden Eggs and Hyperbolic Discounting, Quarterly Journal of Economics, May 1997: 443-477.
Lakoff, George (2008): The Neural Theory of Metaphor, in: Gibbs (2008): 17-38.
Landreth, Anthony / Bickle, John (2008): Neuroeconomics, Neurophysiology and the Common Currency Hypothesis, Economics & Philosophy 24(3): 419-429.
Lawson, Tony (1997): Economics and Reality, London: Routledge.
Lea, Stephen G. / Webley, Paul (2006): Money As Tool, Money As Drug: The Biological Psychology of a Strong Incentive, Behavioral and the Brain Sciences 29: 161-209.
LeDoux, Joseph (1997): The Emotional Brain. The Mysterious Underpinnings of Emotional Life, New York: Simon & Schuster.
Lewontin, Richard (2007): The Genotype/Phenotype Distinction, The Stanford Encyclopedia of Philosophy (Spring 2007 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2007/entries/genotype-phenotype/>.
Macdonald, Graham / Papineau, David, eds. (2006b): Teleosemantics. New Philosophical Essays, Oxford / New York: Oxford University Press.
Maynard Smith, John (2000): The Concept of Information in Biology,
Philosophy of Science 67(2): 177-194.
McCabe, Kevin A. (2008): Neuroeconomics and the Economic Sciences, Economics & Philosophy 24: 345-368.
Menary, Richard (2007): Writing as Thinking, Language Sciences 29: 621-632.
Menger, Carl (1892): On the Origins of Money, Economic Journal 2: 239-55. http://socserv.mcmaster.ca/econ/ugcm/3ll3/menger/money.txt

Mesoudi, Alex (2011): Culture and the Darwinian Renaissance in the Social Sciences and Humanities, Journal of Evolutionary Psychology 2011: 1-16, DOI: 10.1556/JEP.9.2011.29.1.
Mesoudi, Alex / Whiten, Andrew / Laland, Kevin N. (2006): Towards a Unified Science of Cultural Evolution, Behavioral and the Brain Sciences 29: 323-383
Millikan, Ruth (1989): Biosemantics, Journal of Philosophy 86: 281-297.
Millikan, Ruth (2005): Language: A Biological Model, Oxford: Clarendon.
Millikan, Ruth G. (2009): Biosemantics, in: McLaughlin et al. (2009): 394-406.
Neander, Karen, "Teleological Theories of Mental Content", *The Stanford Encyclopedia of Philosophy (Winter 2009 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2009/entries/content-teleological/>.
Noë, Ronald / van Hooff, Jan A.R.A.M. / Hammerstein, Peter, eds. (2001): Economics in Nature. Social Dilemmas, Mate Choice and Biological Markets, Cambridge et al.: Cambridge University Press.
North, Douglass C. (2005): Understanding the Process of Economic Change, Princeton and Oxford: Princeton University Press.
Odling Smee, F. John / Laland, Kevin N. / Feldman, Marcus W. (2003): Niche Construction. The Neglected Process in Evolution, Princeton:Princeton University Press.
Ofek, Haim (2001): Second Nature. Economic Origins of Human Evolution, Cambridge: Cambridge University Press.
Oullier, Olivier / Kelso, J.A. Scott (2009): Social Coordination, from the Perspective of Coordination Dynamics, in: Encyclopedia of Complexity and Systems Sciences, ed. R.A. Meyers, Berlin: Springer.
Oullier, Olivier / Kirman, Alan P. / Kelso, J.A. Scott (2008): The Coordination Dynamics of Economic Decision Making: A Multilevel Approach to Social Neuroeconomics, IEEE Transactions on Neural Systems and Rehabilitation Engineering 16(6): 557-571.
Oyama, Susan (2001): The Ontogeny of Information. Developmental Systems and Evolution. Durham: Duke University Press.
Papineau, David (2007): Naturalism, The Stanford Encyclopedia of Philosophy (Spring 2007 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2007/entries/naturalism/>.
Papineau, David (2009): The Causal Closure of the Physical and Naturalism, McLaughlin et al. (2009): 53-65.
Pecher, Diane / Zwaan, Rolf A., eds. (2005): Grounding Cognition. The Role of Perception and Action in Memory, Language and Thinking, Cambridge: Cambridge University Press.
Pham, Michael Tuan (2004): The Logic of Feeling, Journal of Consumer Psychology 14(4): 360-369.
Pham, Michel Tuan (2007): Emotion and Rationality: A Critical Review and Interpretation of Empirical Evidence, Review of General Psychology 11(2): 155-176.
Phelps, Elisabeth A. (2009): The Study of Emotion in Neuroeconomics, in: Glimcher et al. (2009): 233-250.
Pinch, Trevor / Swedberg, Richard, eds. (2008): Living in a Material World. Economic Sociology Meets Science and Technology Studies, Cambridge: MIT Press.
Pinker, Steven (2007): The Stuff of Thought. Language as a Window into Human Nature, New York: Penguin.
Preda, Alex (2005): The Investor as a Cultural Figure of Global Capitalism, in: Knorr Cetina / Preda (2005): 141-162.
Preda, Alex (2008): STS and Social Studies on Finance, in: Hackett et al. (2008): 901-920.
Prinz, Jesse (2009): Emotions: Motivating Feelings, in: McLaughlin et al. (2009): 678-690.
Pryor, Frederic L. (1977): The Origins of the Economy. A Comparative Study of Distribution in Primitive and Peasant Societies, New York et al.: Academic Press.

Radnitzky, Gerard / Bartley III, W.W., eds. (1987): Evolutionary Epistemology, Rationality, and the Sociology of Knowledge, La Salle: Open Court.
Redish, A David / Jensen, Steve / Johnson, Adam (2008): A Unified Framework for Addiction: Vulnerabilities in the Decision Process, Behavioral and Brain Sciences 31: 415-487.
Reimann, Martin / Bechara, Antoine (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.
Richerson, Peter J.and Boyd, Robert (2005): Not By Genes Alone. How Culture Transformed Human Evolution. Chicago: University of Chicago Press.
Ross, Don / Sharp, Carla, Vuchinich, Rudy E. / Spurrett, David (2008): Midbrain Mutiny. The Picoeconomics and Neuroeconomics of Disordered Gambling, Cambridge, Mass.: MIT Press.
Saad, Gad (2007): The Evolutionary Bases of Consumption, New York: Lawrence Erlbaum.
Santos, Laurie R. / Chen, M. Keith (2009): The Evolution of Rational and Irrational Economic Behavior: Evidence and Insight from a Non-human Primate Species, in: Glimcher et al. (2009): 81-94.
Sapienza,Paola / Zingales, Luigi / Maestripieri, Dario (2009): Gender differences in financial risk aversion and career choices are affected by testosterone, August 2009, PNAS, doi:10.1073/pnas.0907352106
Schantz, Richard, ed. (2004): The Externalist Challenge, Berlin: de Gruyter.
Schlicht, Ekkehard (1998): On Costum in the Economy, Oxford: Clarendon.
Schlicht, Ekkehard (1998): On Costum in the Economy, Oxford: Clarendon.
Schotter, Andrew (1982): The Economic Theory of Social Institutions, Cambridge et al.: Cambridge University Press.
Schultz, Wolfram (2009): Midbrain Dopamine Neurons: A Retina of the Reward System? in: Glimcher et al. (2009): 323-329.
Seabright, Paul (2004): The Company of Strangers: A Natural History of Economic Life, Princeton NJ: Princeton University Press.
Searle, John R. (1995): The Construction of Social Reality, New York: Free Press.
Searle, John R. (2004): Biological Naturalism, http://socrates.berkeley.edu/~jsearle/articles.html
Searle, John R. (2005): What Is an Institution?, in: Journal of Institutional Economics 1(1): 1-22.
Searle, John R. (2010): Making the Social World. The Structure of Human Civilization, Oxford: Oxford University Press.
Shafir, Edward / Diamond, Peter / Tversky, Amos (1997): Money Illusion, Quarterly Journal of Economics CXII(2): 341-374.
Skyrms, Brian (2004): The Stag Hunt and the Evolution of Social Structure, Cambridge et al.: Cambridge University Press.
Skyrms, Brian (2010): Signals. Evolution, Learning, & Information, Oxford: Oxford University Press.
Smit, J.P. / Buekens, Filip / du Plessis, Stan (2011): What is Money? An Alternative to Searle’s Institutional Facts, Economics & Philosophy 27(1): 1-22.
Sperber, Dan (2011): A naturalistic ontology for mechanistic explanations in the social sciences, in: Pierre Demeulenaere (ed.) Analytical sociology and social mechanisms (Cambridge: Cambridge University Press), pp. 64-77.
Sperber, Dan and Hirschfeld, Lawrence A. (2004): The Cognitive Foundations of Cultural Stability and Diversity, TRENDS in Cognitive Science 8(1): 40-46.
Sterelny, Kim (2004): Externalism, Epistemic Artefacts and The Extended Mind, in: Schantz (2004): 239-254.
Strauss, Claudia/Quinn, Naomi (1997): A Cognitive Theory of Cultural Meaning, Cambridge et al.: Cambridge University Press.
Subrahmanyam, Avanidhar (2007): Behavioral Finance: A Review and Synthesis, European Financial Management 14(1): 12-29.

Tooby, John / Cosmides, Leda (2005): Conceptual Foundations of Evolutionary Psychology, in: Buss (2005): 5-67.
Toronchuk, Judith A. / Ellis, George F.R. (2005): Affective Neuronal Darwinism: The Nature of the Primary Emotional Systems, http://web.uct.ac.za/depts/adult-ed/archive/ss2005/papers/AND%20II.pdf
Trepel, Christopher et al. (2005): Prospect Theory on the Brain? Toward a Cognitive Neuroscience of Decision Under Risk, Cognitive Brain Research 23: 34-50.
Trivers, Robert (1985): Social Evolution. Menlo Park: Benjamin/Cummings.
Tuomela, Raimo (1995): The Importance of Us. A Philosophical Study of Basic Social Notions. Stanford: Stanford University Press
Tuomela, Raimo (2007): The Philosophy of Sociality, Oxford: Oxford University Press.
Tuomela, Raimo (2011): Searle’s New Construction of Social Reality, Analysis Reviews 71(4): 707-719.
Turner, Mark (2008): Frame Blending, in: Favretti (2008): 13-32.
Van der Sar, Nico L. (2004): Behavioral Finance: How Matters Stand, Journal of Economic Psychology 25: 425-444.
Veblen, Thorstein (1899/1965): The Theory of the Leisure Class, New York: Kelley.
Wargo, Donald T. / Baglini, Norman A. / Nelson, Katherine A. (2010): The New Millenium’s First Global Financial Crisis: The Neuroeconomics of Greed, Self-interest, Deception, False Trust, Overconfidence and Risk Perception, in: Stanton et al. (2010): 78-100.
Weber, Bernd / Rangel, Antonio / Wibral, Mathhias / Falk, Armin (2009): The Medial Prefrontal Cortex Exhibits Money Illusion, Proceedings of the National Academy of Sciences 106(13): 5025-5028.
West, Robert (2006) Theory of Addiction, Oxford: Blackwell and Addiction Press.
Witt, Ulrich (2003): The Evolving Economy. Essays on the Evolutionary Approach to Economics, Cheltenham and Northampton: Edward Elgar.
Young, H. Peyton (1998): Individual Strategy and Social Structure, Princeton: Princeton University Press.
Zahavi, Amotz / Zahavi, Avishag (1997): The Handicap Principle. A Missing Piece of Darwin’s Puzzle, New York / Oxford: Oxford University Press.
Zaloom, Caitlin (2003): Ambiguous Numbers: Trading Technologies and Interpretation in Financial Markets. American Ethnologist 30(2): 258-272.
Zaloom, Caitlin (2004): The Productive Life of Risk, Current Anthropology 19(3): 365-391.