Schmollers Jahrbuch 136 (2016), 227 – 236
Duncker & Humblot, Berlin

Performativity, Identity and Economic Naturalism: A Comment on John Davis' "Economics, Neuroeconomics, and the Problem of Identity"

By Carsten Herrmann-Pillath*

Abstract

This paper is an extensive comment on John Davis' paper "Economics, Neuroeconomics, and the Problem of Identity" published in this very journal (Vol. 136, No. 1). So far, the methodological assessments of neuroeconomics by economists vacillate between the Scylla of neuro-reductionism and the Charybdis of Friedman's instrumentalism. Following Davis' approach to identity economics, I argue that there is a third way shown by methodological approaches to the neurosciences which are non-reductionist and highlight complex multi-level explanations, including external interactions, such as in distributed cognition theories. I suggest that newly emerging areas of study – such as social neuroscience – can be traced back to G.H Mead's theory of the individual, with important implications for economics.

JEL Codes: B41, D87

Behavioural economics and neuroeconomics play an ambivalent role in contemporary economics. On the one hand, their rise contributes to strengthening the self-acclaimed status of economics as the discipline in the social and human sciences which most closely follows the standards of scientific research as defined by the natural sciences. This is reflected in the fact that important papers in these fields are sometimes published in the flagship journals of the sciences, such as *Nature*. On the other hand, however, many economists continue to argue that these decidedly cross-disciplinary fields have nothing to contribute to the revision of the fundamentals of economic theory. As elucidated in John Davis' paper, this results from the peculiar methodological principles that have guided economic theorizing since the turn to marginalism and subjectivism that was completed in the 1930s. These principles play an essential role in defining

* Max Weber Center for Advanced Cultural and Social Studies, University of Erfurt, PO Box 900221, 99105 Erfurt, Germany. The author can be reached at cahepil@online.de.

economics as a separate and autonomous scientific discipline: Yet, they are often difficult to understand for outsiders, who would mostly expect that claims regarding scientific status also imply that its fundamental principles advance claims about reality. But principles such as rationality and optimization do not permit those claims, as far as their relationship to real-world individuals are concerned. That is why criticism based on insights from other disciplines that investigate real world individual behaviour simply miss the point.

What are the empirical claims of standard economic theory? I think that they exclusively relate to markets, defined as economic systems that have a certain institutional structure that create society-wide or even global systems of computing economic values, in the neoclassical sense of relative scarcities. These systems have an immensely high degree of complexity, and so the interest of economics is both to provide evidence for the possible existence of such systems and, of course, also to show that real-world economic systems are or can be instances of such theoretical markets. If we look at the role of individual behaviour in these systems, we need to recognize that this is endogenous to the operations of the system. Following recent developments in economic methodology, we can say that it becomes *performative*. This can be explained by a number of real-world mechanisms of inducing performative actions, in particular that market systems select agents that display a certain behavioural pattern, or that agents would adopt ideas and theories about the operations of markets that align their behaviour with the requirements of systems sustainability and viability.

Let me provide an example. The seminal case for introducing the notion of performativity to economics is financial markets (MacKenzie 2006). In financial markets we observe a direct interaction between the progress of the theory of finance, the design of market institutions and specific tools that agents deploy, and even many behavioural practices. Tellingly, many decisions in financial markets are today implemented by computers which are directly programmed to perform the system without any disturbances resulting from human psychology. But in fact, in the "good old days" of the pit, traders were agents that underwent (and still do) a long and arduous socialization process in which they learned to play the rules of the game (Zaloom 2004). For example, they acquired practices such as focusing their attention on the present, evidently to avoid the harmful impact of loss aversion. Such a training process is no longer necessary with computers; in fact, good programming suffices. As this example shows, the theoretical markets of finance only exist when they are performed in the real world. This defines a different empirical agenda than either mainstream economists or their critics propose, because this would investigate into the mechanisms of performativity (Herrmann-Pillath 2016a). "Rationality" is not an empirical proposition, but in the first place refers to the conditions that must hold for markets to work efficiently. In this methodological third alternative, performativity becomes the core theoretical notion which makes the real-world

processes explicit as to how these conditions come into existence and are sustained. Consequently, the methodological failure of standard economics only lies in the circumstance that it remains satisfied with naive instrumentalism, as elucidated by John Davis, and that it did not advance to the study of the performative nature of economic systems. Overlooking the latter creates the awkward tension between its theoretical principles and empirical claims.

Indeed, economists were largely responsible for messing up this clear methodological position when they launched imperialist raids into other disciplines, claiming that economics is a "theory of everything" that also reaches to the level of real-world individuals in explaining all kinds of their behavioural phenomena in social domains other than markets. Counter-raids by those disciplines – such as psychology – which regard these phenomena as their domain are perfectly intelligible and also justified. I think that these intellectual battles fail to advance our understanding of economics as a science and are harmful wars of attrition.

The crucial task is to recognize the nature of economics as a discipline that does not study individual behaviour at all but only systems-level phenomena which involve individual actions as constitutive elements. These systems are not "natural" systems because they are being constituted as what John Searle calls "observer relative facts" (1995). That means their existence and emergence depends on cognitive operations of human individuals who interact in groups such that specific forms of collective intentionality arise which result into particular forms of practical rationality. This is the basis of performative actions that work together in creating the systems mostly as unintended effects, in the Hayekian vein of being constituted by human action, but not by human design. Once the systems have emerged, they create different channels by which human action is aligned with systems functioning. The example of financial markets is a case in point: Actors in financial markets assume a form of collective intentionality by which those markets come into existence or are being performed. These actors are as *real* as the individuals of flesh and blood, but they are different. The financial trader is an "observer relative fact" of its own ontological status.

An extremely interesting other instance of this abstract picture is the way how economics assumes the role of actively creating ideational and institutional frames that undergird performative actions in markets. Take the recent movement in advancing "nudging" as a concept for improving individual behaviour: this approach assumes psychology as a starting point, but suggests that there are devices and tools of behavioural design which can change empirically validated behavioural patterns. On surface, we observe a clash between the new field of behavioural economics and so-called neoclassical economics in the first place. But at the same time the concept of nudging reveals that behavioural economics continues to accept the basic claims of standard economics after all. This is because the standard assumptions of rationality serve as the

benchmark towards which nudging aims in re-engineering individual behaviour. For example, there are certain standards concerning the rational management of financial assets, and nudges aim at reducing errors that individuals commit like in the context of pension provisions. Collectively, a system defined by certain institutions of finance, nudges and individual choices would constitute an efficient financial market that also extends to individuals who are not "professional" actors. In spite of the internal critique of standard economics by behavioural economics, the latter, in fact, remains part and parcel of the performative role of the former. As long as the standard of *rationality* is taken as the reference to empirically determine "failures" of individual choices, behavioural economics remains an important part of the universal performative project of economics and does not effectively question standard economics.

We need to observe that most debates about the methodology of economics fail to recognize that economics is endogenous to the economic systems that it claims to describe and analyse, in the sense of having performative functions. This observation resolves the intricate difficulties that have always emerged from mixing up positive and normative criteria in economic research. The notion of performativity goes beyond this juxtaposition and thus explains this factual mix in economics as it is practised. This is because performativity is not simply based on normative criteria by which a certain behaviour is imposed on individuals: performativity results from changing the ways how individuals collectively perceive the world and from the resulting actions that work together in creating a particular social ontology, which then appears to be manifest in empirical phenomena conducive to positive research. This is why in spite of the many criticisms by behavioural finance, the established theory of finance remains at the core of the discipline.

In modern cognitive sciences, we can restate this insight in recognizing the nature of markets as systems of distributed cognition (Hutchins 1995; Clark 2011). This is an idea that, using a different language, was already advanced by Hayek (1945) in his notion of markets as forms of knowledge dispersed in society which cannot be centralized in any single entity: Hayek famously argued that the price system is an information-processing system that coordinates individual actions without presupposing that individuals themselves have perfect and complete information about the conditions of the economy and even their individual options of choice. I think that one of the most exciting venues of future research is showing how behavioural economics and neuroeconomics lead us to approach this phenomenon from the individual level. John Davis points the way in arguing that neuroeconomics in fact investigates systems of distributed cognition, given its underlying empirical standards and methodologies. Interestingly, Hayek's thinking about markets was deeply influenced by his early work on the human brain as a "sensory order" (Hayek 1952). I argue that the further we push the boundaries of economic research to the level of neuronal phenomena, the more we will be driven towards recognizing the dis-

tributed cognition function of markets, which includes the endogenous evolution of economics as being part and parcel of these functions. As a result, I advance the claim that radicalizing the *naturalization* of economics via neuroeconomics will end up with recognizing the distributed cognition function of markets and the phenomenon of performativity that stands at its core.

How is that possible? I expect that once Pandora's box of cross-disciplinary integration is opened, economics has to accept a methodological integration across those different disciplines. In such an integration, certainly the position of naive instrumentalism has to be given up in order to align economics with the standards of the sciences (de facto, at least, though probably not in the rhetoric of inter-disciplinary intellectual battles). Now, it is interesting to notice that instrumentalism in economics has always been wedded with an implicit allegiance to the "covering law" approach to scientific disciplines, which may be the main culprit why even economists themselves often misunderstand the empirical status of their fundamental analytical principles. In comparison, the life sciences and the neurosciences, in particular, do not adopt this methodological principle but aim at making so-called "constitutive explanations" operational (Craver 2007). The neurosciences investigate complex mechanisms by which certain empirical phenomena are generated, which in turn reflect more general laws of nature; but those laws are by far too unspecific to achieve valid explanations, especially if we also have specific action in mind, such as applied therapeutic use of the neurosciences. This approach of constitutive explanations has also recently gained acceptance in the other social sciences (Demeulenaere 2011; Craver and Tabery 2016).

Mechanisms are complex multi-level phenomena with generative powers. In the neurosciences, this means that it is impossible to reduce behavioural phenomena to the level of simple neuronal processes having physical and chemical basic properties. In other words, doing causal analysis in dissecting mechanisms does not imply reductionism. This is not the place to go into details, so let me just state that a fundamental property of complex mechanisms in the neurosciences is contextualization: single and simple causal chains with a neurophysiological substrate are always contextualized via the simultaneous interactions with the environment in which these chains materialize. The complexity of mechanisms means that this environment is multi-faceted, relating a single causal chain to the immediate environment of other neuronal connections in which it happens, resulting in feedback loops that shape the causal chain in turn, but also going further in connecting the part of the brain in which the causal chain operates with other parts of the brain. Finally, this includes the interaction with the external environment, especially as being mediated by symbolic structures that are a defining feature of human cognition and action, such as language.

Once we combine economics and neurosciences in the emerging field of neuroeconomics, the integration needs to be methodologically grounded on the

principles of constitutive explanations (Herrmann-Pillath 2016b; compare Ross 2012). This implies that reductionism à la Glimcher (2011) is not a feasible option, i.e. claiming that foundational concepts such as utility could be reduced to neurophysiological phenomena, thus positing what Camerer (2012) has aptly called a "neuroclassical" theory. This view overlooks the fundamental fact that precisely if we move to the level of neuronal phenomena, these have to be put into the context of complex multi-level mechanisms of distributed cognition. If economists want to explain observed behaviour, they will have to analyse those mechanisms, while taking the intermediating causal role of neurophysiological phenomena into consideration. Obviously, this would lead towards the inclusion of higher-level cognitive functions in the full picture. These functions are in turn mediated not only by neurophysiological mechanisms, but by a vast array of externalized "scaffolds," which include, in particular, language, cultural symbolic media, and regularized patterns of action in human groups (cf. North 2005). So far, these scaffolds are mostly the object of research in the other social sciences, such as sociology or anthropology, but also in the humanities. Hence, I predict that it is progress in the neurosciences that ultimately undergirds the re-integration of economics with the other disciplines in the human sciences (cf., for example, Damasio 2010).

I also think that there are different venues in which we can achieve such a methodological re-integration. One is to go back to the roots and consider the intellectual history of the current disciplinary segregation. Quite fascinating, the explicit discussion of the sciences and science-based psychology played an important role in driving the separation of the modern social sciences and the humanities (as an important example, see Dilthey 1883). This debate goes back to Hegel in his famous critique of what was the equivalent to the neurosciences in his days: phrenology (Herrmann-Pillath and Boldyrev 2014). Hegel influenced many thinkers throughout the 19th and the early 20th century. One author who is particularly relevant in the current context is George Herbert Mead, who is recognized as the father of the discipline of "social psychology." Interestingly, today we see the emergence of a sub-discipline of social neuroscience. Mead (1934) argued that the brain (he mostly used the term 'central nervous system') is essential for mediating human action in social contexts, and he explicitly referred to a biological reference frame, especially Darwinian theory. However, just because he did so, he approached human action in terms of its fundamental achievement in adapting to the natural environment. This achievement is cooperation. In human cooperation, the core capacity is using language for coordinating action. For Mead language is the essential analytical template by which human behaviour has to be approached (for a contemporary related view, see Tomasello 2008).

With hindsight, we can tell that Mead anticipated the views in modern neuroscience which approach the human brain as being a "social brain" (Frith 2007). This also implies that for developing a workable brain, social interaction

is essential; what is reflected is the fact of human neoteny. That means, we cannot just ask how the brain embodies the capacity for language: vice versa, being immersed in an environment replete with linguistic phenomena is essential for developing a human brain to biological maturity. This was already recognized by Mead: for him, brains do not have minds, but minds involve brains. A mind is a phenomenon that is socially constituted, in particular via the mediation of language.

The implication for economics is straightforward: Mead's analysis, as reinstated by modern neuroscience and parts of neurophilosophy, implies that the notion of an *agent* is complex and cannot be reduced to the *individual* as demarcated by the boundaries of the body. Agents are distributed in social interaction – and in communication in particular. This view matches with the modern approaches to distributed cognition, so that we can also speak of "distributed agency:" the agent is not a given that is determined by a bodily counterpart, but emerges in sociality, and is socially assigned to the body, as assumed in modern Western society. Interestingly, this is exactly what the previous analysis of performativity implies, and we can reach a surprising conclusion. An "individual," in the sense of economics, may exist, but only as a manifestation of peculiar mechanisms of distributed agency. This is what already transpired in the previous analysis of financial traders and markets: a financial trader is a distributed agent that emerges from the interaction between organismic features and external scaffolds such as exchange technologies, symbolic media and so forth. This distributed agent is conventionally assigned to the physical body of the trader, but that does not mean that the latter is ontologically identical with the former.

Mead caught this complex phenomenon with his famous distinction of the "I" and the "me" in his theory of the self, which is in turn anchored in his sophisticated discussion of imitation (again, compare Tomasello 2008). In this, he anticipated the modern debates about the role of mirror neurons in human action (Fogassi 2011). Considering the various approaches to multiple selves in modern economics also mentioned by John Davis, Mead proposed a distinct and very interesting alternative. According to him, human imitation is unique, since mediation by symbolic media enables a capacity to adopt the attitude of others to one's own actions, thus also enabling anticipatory action. This kind of complex circular interaction constitutes the "me" as one essential part of individual agency. The "me" is multiple, as it enables role-taking as a distinctive feature of human behaviour. This, in my view, is exactly what makes performativity possible.

Interestingly, the "me" seems to be the phenomenon that is dealt with explicitly by recent economic theories of identity that strive to keep as close as possible within the disciplinary confines of economics (Akerlof and Kranton 2000). But as John Davis (2007) has shown, these theories only result in a picture of human beings as assemblies of social roles, hence it is lacking a convincing account of personal identity. This clearly manifests the general weak-

ness of the economic framework in presenting convincing accounts of identity. Mead was already one step ahead since he posited the interaction between "I" and "me" in explaining the self.

The "I" is a concept that is difficult to understand on first sight, especially with an economic background (See Meade 1934, 173ff and 196 and 277). This is because Mead is very insistent in keeping his general argument focused on the social nature of the mind and the self, which is quite different from the economic approach to identity which remains wedded to the concept of utility function in the context of game theory. That means, the different social roles of the "me" would be subject to the maximization of utility assigned to them by the rational agent involved in strategic interactions. Mead's "I" is not a rational agent but a constitutive part of the self that consists of both "I" and "me" and which represents the original creativity of the person relative to internalized social roles (cf. Joas 1992). The "I" is the response of the self to the "me," hence we might envisage an internal dialogic relationship. Now, it is of utmost importance to recognize that the "I" is only accessible to the self via the observation of the resulting actions: the "I" is part of consciousness only via memory, even in the shortest term. We experience our own actions as observers whose perception is already symbolically mediated, though internalized. As such, the "I" is immediately manifest in a new metamorphosis of the "me," as the observed actions are cognitively processed via symbolic intermediation. In other words, the "I" is the flow of creative actions that result (in modern terminology) in an autobiographical narrative, or the "I" is a principle of creativity that makes up the uniqueness of the human individual and that is embodied in narratives of personal development which are part and parcel of the "me" (for related views in the neurosciences, see Damasio 2010; cf. Ross 2007).

I think that Mead's approach is a highly attractive option that differs from all other alternatives to integrating the neurosciences, psychology and economics offered so far. The weakness of these other approaches lies in the fact that they remain reductionist in substituting the economic model of rationality with an equally individualistic, though naturalistic account of behaviour that traces behaviour back to other internal determinants of action, such as distinguishing between different *systems* of decision-making in the brain (e.g. "conscious" versus "habitual," or "slow" and "fast," etc.; Camerer et al. 2005; Kahneman 2011). This view overlooks the essential role of sociality in constituting the human agent. At the same time, there is no room for understanding the core phenomenon of human individuality – the creativity of the person. But we can even provide a neuroscientific account of this if we approach the brain as a highly complex system that is materially unique for every individual, and which at the same time constantly generates novelties in terms of evolving neurophysiological processes and outputs (Edelman 2006, 98ff). Hence, we might speculate that the Meadian "I" is actually the creative brain, and that for a human person emerging from and stabilizing in this creative dynamics, the "me" is essential.

Summarizing, I think that naturalizing economics by moving into the field of neuroeconomics will result in recognizing the *second nature* of our human existence and will contribute to a genuine cross-disciplinary synthesis that not only integrates economics with the sciences, but will also strengthen the analytical significance of the other social sciences and the humanities in understanding human action. Alas, economists so far only realize one side of this cross-disciplinary integration, because they erroneously interpret naturalization as strengthening reductionist claims in economics, and because they stick to instrumentalism as a methodological beacon, as John Davis aptly points out, thus creating a methodological mess. A proper reading of leading neuroscientists such as Damasio (2010), Edelman (2005) or Singer (2002) clearly reveals this delusion of scientism.

References

Akerlof, G. A. and R. E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* CXV (3): 715 – 753.

Camerer, C. 2013. "A Review Essay about Foundations of Neuroeconomic Analysis by Paul Glimcher." *Journal of Economic Literature* 51 (4): 1155 – 1184.

Camerer, C., G. Loewenstein and D. Prelec. 2005. "Neuroeconomics: How Neuroscience Can Inform Economics*." Journal of Economic Literature* XLIII: 9 – 64.

Clark, A. 2011. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.

Craver, C. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of the Neurosciences.* New York: Oxford University Press.

Craver, C. and J. Tabery. 2015. "Mechanisms in Science." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. <http://plato.stanford.edu/archives/win 2015/ entries/science-mechanisms/>.

Damasio, A. R. 2010. *Self Comes to Mind: Constructing the Conscious Brain*. New York: Pantheon.

Davis, J. B. 2007. "Akerlof and Kranton on Identity in Economics: Inverting the Analysis." *Cambridge Journal of Economics* 31(3): 349 – 362.

Demeulenaere, P., ed. 2011. *Analytical Sociology and Social Mechanisms*. Cambridge: Cambridge University Press.

Dilthey, W. 1883. *Einleitung in die Geisteswissenschaften. Versuch einer Grundlegung für das Studium der Gesellschaft und der Geschichte*. Vol. 1. Leipzig: Duncker & Humblot. <http://www.deutschestextarchiv.de/dilthey_geisteswissenschaften_ 1883/ 10>.

Fogassi, L. 2011. "The mirror neuron system: How cognitive functions emerge from motor organization." *Journal of Economic Behavior & Organization* 77(1): 66 – 75.

Frith, C. 2007. "The Social Brain?" *Philosophical Transactions of the Royal Society B* 362: 671–678.

Glimcher, P. W. 2011. *Foundations of Neuroeconomic Analysis*. Oxford and New York: Oxford University Press.

Hayek, F. A. 1945. "The Use of Knowledge in Society." *American Economic Review* XXXV(4), 519–530.

__ 1952. *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology.* Chicago: University of Chicago Press.

Herrmann-Pillath, C. 2016a. "Performative Mechanisms." In *Enacting Dismal Science: New Perspectives on the Performativity of Economics*, edited by Ivan Boldyrev and Ekaterina Svetlova, 53–86. Palgrave MacMillan.

__ 2016b. "Constitutive Explanations in Neuroeconomics: Principles and a Case Study on Money." Forthcoming in *Journal of Economic Methodology.*

Herrmann-Pillath, C. and I. Boldyrev. 2014. *Hegel, Institutions and Economics: Performing the Social*. Abingdon and London: Routledge.

Hutchins, E. 1995. *Cognition in the Wild*. Cambridge and London: MIT Press.

Joas, H. 1992. *Die Kreativität des Handelns*. Frankfurt: Suhrkamp.

Kahneman, D. 2011. *Thinking, Fast and Slow.* London: Penguin.

MacKenzie, D. 2006. *An Engine, Not a Camera: How Financial Models Shape Markets.* Cambridge and London: MIT Press.

Mead, G. H. 2015 [1934]. *Mind, Self & Society.* Chicago and London: Chicago University Press.

North, D. C. 2005. *Understanding the Process of Economic Change*. Princeton and Oxford: Princeton University Press.

Ross, D. 2007. "H. sapiens as Ecologically Special: What Does Language Contribute?" *Language Studies* 29: 710–731.

__ 2012. "Neuroeconomics and Economic Methodology." In *The Elgar Companion to Recent Economic Methodology*, edited by John B. Davis and D. Wade Hands, 61–93. Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Searle, J. R. 1995. *The Construction of Social Reality.* New York: Free Press.

Singer, W. 2002. *Der Beobachter im Gehirn: Essays zur Hirnforschung*, Frankfurt: Suhrkamp.

Tomasello, M. 2008. *Origins of Human Communication*. Cambridge and London: MIT Press.

Zaloom, C. 2004. "The Productive Life of Risk." *Current Anthropology* 19 (3): 365–391.