Article

Theory Culture & Society

Theory, Culture & Society
0(0) 1–27

DOI: 10.1177/0263276414551996
tcs.sagepub.com

SAGE

Critical Neuroscience and Socially Extended Minds

Jan Slaby
Freie Universität Berlin

Shaun Gallagher
University of Memphis

Abstract

The concept of a *socially extended mind* suggests that our cognitive processes are extended not simply by the various tools and technologies we use, but by other minds in our intersubjective interactions and, more systematically, by institutions that, like tools and technologies, enable and sometimes constitute our cognitive processes. In this article we explore the potential of this concept to facilitate the development of a critical neuroscience. We explicate the concept of cognitive institution and suggest that science itself is a good example. Science, through various practices and rules, shapes our cognitive activity so as to constitute a certain type of knowledge, packaged with relevant skills and techniques. To develop this example, we focus on neuroscience, its cultural impact, and the various institutional entanglements that complicate its influence on reframing conceptions of self and subjectivity, and on defining what questions count as important and what kind of answers will be valued.

Keywords

cognitive institution, critical neuroscience, critique, extended mind, social interaction

The concept of the extended mind (Clark, 2008; Clark and Chalmers, 1998) has been a productive but also provocative way of thinking about how various tools and technologies enter into cognition. If we think of these extra-neural processes – our use of instruments for cognitive purposes, but also our bodily movements in gestures, for example – as epistemic acts constitutive of (and not just causally contributing to) cognition, we also need to think of the mind and brain in a rather unorthodox way. Indeed, the concept of an 'extended' mind may be

Corresponding author: Jan Slaby. Email: jan.slaby@fu-berlin.de
Extra material: http://theoryculturesociety.org/

better stated not as something that extends from the brain outward to encompass various tools, constrained by what Clark and Chalmers call the 'parity principle',[1] but as a larger (or *extensive*) process that includes neural processes as one necessary but not sufficient component (see Hutto and Myin, 2013). This not only moves us beyond the standard Cartesian idea that cognition is something that happens in a private mental space, 'in the head', but also towards a more enactive and intersubjective conception of the mind.

We note that there are ongoing debates between enactive and extended theories of cognition. The enactive view (originating with Varela et al., 1991) takes cognition to be more radically embodied than the functionalist version of extended mind defended by Clark (see e.g. Thompson, 2007). It also tends to be anti-representationalist (Hutto and Myin, 2013), in contrast to representational or even minimally representational conceptions found in Clark (2008), Rowlands (2009), Wheeler (2010) and others. Enactive approaches also offer a characterization of the precise coupling that links brain, body, and environment which is different from the supervenience model that subtends the extended mind (see Di Paolo, 2009). Despite these differences, several theorists (e.g. Gallagher and Miyahara, 2012; Menary, 2007) argue that only by integrating a more embodied, enactive and more fully action-oriented concept of intentionality will the extended mind hypothesis be able to fend off the various criticisms leveled against it by internalists like Adams and Aizawa (2008, 2010), and Rupert (2009).

Consistent with this enactive or integrative approach, but also taking things in an even more extensive direction, the concept of a *socially* extended mind (Gallagher, 2013; Gallagher and Crisafi, 2009) suggests that our cognitive processes are extended not simply by the various tools and technologies we use, but by other minds in our intersubjective interactions, and more systematically when we couple with institutions that, like tools and technologies, help to constitute our cognitive processes. So-called 'mental institutions' (Gallagher and Crisafi, 2009), or as we'll call them, 'cognitive institutions', consist of those practices, rules and structures that have been instituted for cognitive purposes (such as making judgments, making decisions and solving problems) in previous activities that are both cognitive and social. It has also been suggested that the notions of socially extended minds and cognitive institutions can facilitate a critical perspective on what such practices, rules and structures do to our cognitive processes (Gallagher, 2013).

In this article we explore the potential of this concept of the socially extended mind to facilitate the development of a critical neuroscience. We first explicate the concept of cognitive institution and show how it builds on a more enactive version of the extended mind. One clear example of a cognitive institution, the legal system, has been the focus of previous analysis (Gallagher, 2013), and we will briefly rehearse that

analysis here. We then turn to the idea that science itself is a good example of an established cognitive institution that, through various practices and rules, shapes our cognitive activity so as to constitute a certain type of knowledge, packaged with relevant skills and techniques. Building on this idea, we focus on neuroscience. Our intent is to show that, by understanding neuroscience as a cognitive institution – that is, as a set of practices that help us to think about and solve problems within a specific domain – we gain a critical perspective on what neuroscience accomplishes. As such, this is an essay in critical neuroscience.

Cognitive Institutions

Cognitive institutions are institutions that help us to accomplish certain cognitive tasks, and they do so in a way that contributes to the constitution of the cognitive process. Indeed, without them, specific classes of cognitive processes would simply not exist. Examples include things like legal systems, educational systems, cultural institutions like museums, and even the institution of science itself. In each case a cognitive institution:

1. includes cognitive practices that are produced in specific times and places, and
2. is activated in ways that extend our cognitive processes when we engage with them (that is, when we interact with or are coupled to these systems in the right way).

The concept of coupling, which defines the way that we engage with such systems, is important for establishing that this is a matter of constitution rather than just a causal relation. On a more standard view of the mind, all of the work is done in the brain; cognition supervenes exclusively on neuronal processing. Only neurons, as vehicles of cognition, can constitute genuine cognition. The use of a tool or technology, that is, the establishment of an instrumental relation with such an externality, then, would simply facilitate (and enter as a causal factor in) a process of cognition. If the use of a tool or technology, however, defines a new dynamic system without which the cognitive process would not be what it is, that is, if the way that we are coupled to the tool or technology is such that this coupling actually makes the cognition the kind of thing that it is, then we consider this to be a constitutive relation (see De Jaegher et al., 2010 for more on this topic).

What coupling is in any particular instance may vary, depending on the nature of the externality (artifact, tool, technology or institution) but the key aspects of enactive coupling are (1) that it is a dynamic process (i.e. one in which a co-dependence is established between the coupled systems such that what happens in or to one system is partly dependent on the situation of the other); (2) that the recurrent engagement with the externality leads to a structural congruence (Thompson, 2007: 45);

(3) that the engaging organism (or agent) maintains its autonomy (its own internal self-organization).[2]

The legal system is a good example of a cognitive institution. To see this one needs to consider specific practices within the legal system, rather than considering the legal system as an abstraction. For example, a contract is an expression (in this case a legal agreement) of several minds, establishing in external memory an agreed-upon decision, adding to a system of rights and laws that transcend the particularities of any individual's mind. Contracts are cognitive products that, in turn, contribute to and shape our cognitive processes in further thinking or problem solving. Institutions of property, contracts and rights not only guide our thinking about social arrangements, for example, or about what we can and cannot do, but also allow us to think in ways that are not possible without such institutions. Insofar as we cognitively engage with such instruments and institutions we extend and transform our cognitive processes.

The legal system is constructed in part in these cognitive processes, and the uses of the legal system in the administration of justice, or the application of law to particular cases, are cognitive processes. These are not processes, however, that happen simply in the individual brains of judge, jury, defense attorney, prosecutor, etc. In a court of law, for example, evidence and testimony are produced and judgments are made following a set of rules that are established by the system. The process in which the judgments get made will depend on a body of law, the relevant parts of which come to the fore because of the precise particulars of the case, and as we remain cognitively engaged in the proceedings. The procedures, practices and rules have been instituted in previous cognitive practices and they create the tracks along which ongoing cognitive processes must run to keep them, literally, legitimate.

On this view, the kinds of judgments that count as legal judgments are not confined to individual brains, or even to the plurality of brains that constitute a particular court. They emerge in the workings of a large and complex set of pre-defined practices and are cognitive processes that further contribute to the continued working of the system in the form of precedents. A judgment made in such contexts is a form of cognition that comes about only because a number of people are properly coupled to (dynamically engaged with) a large and complex system without which the judgment could not happen. Indeed, as a case of legal judgment per se, it is a cognitive practice that, in principle, could not happen just in an individual head. For an individual to make such judgments he or she must interact with the legal institution to form a structurally coupled system in a way that allows new cognitive processes to emerge. Take away the external part of this cognitive process – take away the legal institution – and 'the system's behavioural competence will drop, just as it would if we removed part of its brain' (Clark and Chalmers, 1998: 9).

More generally, any such institution is created by means of our own (shared) mental processes, or we inherit it as a product constituted in mental processes already accomplished by others. When we engage with these institutions for purposes of further cognitive work, and thereby constitute a case of complex epistemic action, we establish a dynamic process that allows us to solve problems or control behavior and action. Such institutions allow us to engage in cognitive activities that we are unable to do purely in the head, or even in many heads; they make this a particular kind of cognition (e.g. a legal decision that has a specific kind of effect) that it would otherwise not be. A legal system is only one example of this kind of extended sense-making process; other types of institutions, including political, military, economic, religious and cultural institutions, as well as science itself, are further examples.

Neuroscience as a Cognitive Institution

The picture just painted is somewhat oversimplified since the legal system does not function in isolation from other cognitive institutions. Generally, this can be said of any such institution. Consider, for example, the cognitive work involved in scientific research. Would it be possible – or would science be what it is – without the kinds of things – labs, instruments, scientific practices and procedures (including publication practices that specify stylized sections in the standard journal article) – that carry scientific thinking along and that allow for the practice of science? It is possible to list a large set of institutional practices that make instances of cognition scientific cognition. This would be a study in itself. To be clear, however, we know that research questions and decisions in science are not determined purely by scientific procedure, and scientific results are not strictly confined to scientific labs. Accordingly, to fully understand the ins and outs of scientific practices one must consider not only the institution of science but also how precisely this institution is related to other cognitive institutions and how it is embedded within the ambient society at large. Clearly, this is a vast undertaking. To circumscribe our subject matter, we focus on some issues that pertain to human-level neuroscience, as a particular set of scientific practices currently much discussed, and on how such practices relate to other kinds of institutions, such as medical, media-related, legal and broader cultural institutions.

In this regard we intend to discuss the role played by neuroscience with respect to issues such as biomedical thinking, but also, and especially, with respect to ongoing attempts to understand human beings (human behavior, subjectivity and selfhood), and how such conceptions affect other cognitive institutions, such as the legal system. Moreover, we illustrate some of the ways in which neuroscience itself is pervaded by broader cultural practices and tendencies, such as those pertaining to

the media and, even more generally, to the contemporary capitalist economy.

In doing so, we want to show two things. First, that once we open discussion to include broader cultural factors, the institution we are considering will be more fluid and informal than, for example, the legal system considered on its own. Second, the notion of a cognitive institution is itself a helpful tool for developing a critical stance that allows us to scrutinize current institutional practices. Critique here takes the form of assessments of an institution's modes of operation and de facto impacts, analyzed against the background of its official and unofficial aims, purpose and directions. How does the operational reality of an institution and its specific effectiveness measure up to the ideas and principles that have led to its creation? On a more general level, critique also implies asking whether some given institutional procedures improve (or impede, or distort) our understanding, our communicative practices, our possibilities for action, our recognition of others, our shared and circumscribed freedoms, and so forth.[3]

Given the widespread withdrawal, on the part of cultural theorists and humanities scholars, from more rigorous and more politically ambitious approaches to critique in the past decades – Latour (2004) provides both a reflection on and exemplification of this trend – we believe that our approach can re-open a discussion on possible ways to engage critically with parts of the institutional realities of our contemporary lifeworlds. With this, we oppose those recent socio-cultural analysts of neuroscience who consciously remove 'critique' from their methodological toolbox. In particular, we have in mind Nikolas Rose's and Joelle Abi-Rached's otherwise impressive assessment of the recent cultural ascendency of neuroscience (Rose and Abi-Rached, 2013; see Cooter and Stein [2013] for a positioning somewhat similar to our own).

As neuroscience is a vast field that spans a multitude of different approaches, and even different disciplines ranging from molecular neuroscience and biological psychiatry to systems neuroscience and neuroinformatics, we will restrict our treatment to a select subdomain: the recently burgeoning *social, cognitive and affective neurosciences* – SCAN, in short. The acronym SCAN serves a double purpose, as it not only abbreviates the scientific fields collected under its scope but also highlights the relevance of *brain scans* to this neuroscientific subfield. PET, fMRI and MEG brain imaging have not only led to tremendously productive lines of experimental research, but have also considerably fueled the growing public interest in the neurosciences in recent years. Another reason for the broad appeal of the SCAN disciplines lies in their exquisite selection of research themes: human social cognition, human emotions, decision-making, moral judgment, empathy, consciousness – without exception issues of the highest and most general interest, as they arguably concern the very core of what it is to be human. The SCAN disciplines aspire to no less than becoming a leading scientific approach

to human nature (or rather, to *human nature 2.0*, as anthropologist Allan Young [2012] has quipped insightfully).

We will first sketch some of the central features of the kind of cognitive institution that the SCAN disciplines have become in recent years. We will consider the specific combination of disciplinary procedures and practices, the ways the field deals with the specifics – especially the almost unimaginable complexity – of its subject matter, how it thereby inspires and mobilizes philosophical approaches to its target phenomena, and we will trace select aspects of how it is connected to other social, cultural and cognitive institutions that are relevant in modern societies. By way of highlighting some problematic recent developments, we will motivate a critical perspective that serves the aim of reflectively scrutinizing some of neuroscience's institutional routines and their effects upon society. To establish such reflective practice as part of the operations of science itself – instead of just external commentary and critique – is a central aim of critical neuroscience (see e.g. Choudhury and Slaby, 2012).

Things that Matter

In thinking about cognitive institutions, traditionally inclined philosophers might tend to overstate the importance of abstract, textual or procedural features – such as rules and regulations, laws, codified procedures, protocols, etc. – while neglecting the relevance of 'stuff': the material things that matter to the institutional practice in question. In the case of science, classical cognitivists might still think that ideas, texts, knowledge, skills – apparently possessed or mastered by individual scientists – are closer to the essence of science than the various 'hardwares' operative in scientific work: laboratories, experimental equipment, technologies such as measurement devices or computers, but also meeting rooms, work gear, lab animals, or more diffuse features such as the overall architectural layout of scientific facilities and the specific 'atmosphere' pertaining in these. Materials, buildings and places are the crucial enabling conditions of scientific practices. Machines are objectifications of knowledge and their operation seriously reduces or distributes the cognitive load to be handled and the level of practical skill to be mastered by the individual scientists operating them. Accordingly, current science would not be what it is were it not for the materials it employs, the plethora of physical objects that drive it along, the machines it puts to use, the specific sites in which it is set, the laboratory 'microworlds' it creates, and the manifold skills needed to construct, operate, fix and adjust all the relevant instrumentation (see e.g. Latour, 1999; Pickering, 1995; Rheinberger, 1997; Rouse, 1987). Leading technologies – such as particle accelerators or fMRI scanners – can define entire fields of inquiry, and scientists often spend more time constructing, modifying and calibrating equipment than they spend observing

phenomena, devising theories or writing papers. All of these things together in inextricable, local entanglements – the laws, codified procedures, protocols, embodied and enacted in vast material cultures and technological practices – help to make science the cognitive institution that it is.

In the case of SCAN, neuroscience crucially revolves around the MRI scanner and other sophisticated data-gathering technologies. Indeed, brain scanners epitomize many of the features that are central to the cognitive institution that is today's human-level neuroscience – both its valuable and productive aspects and its potential dangers and distortions. Accordingly, in the following, we will take the MRI scanner as our point of departure, illustrating the issues we want to raise about neuroscience, as a cognitive institution, and its interconnections with other mental institutions, by describing some of the machine's central features and effects.

The Objectivity Machine

The great advantage of the MRI technology is its non-invasive brain measurement procedure. While the basic principle of MRI is based on radio signals emitted by hydrogen nuclei when placed within a strong magnetic field, the basic trick behind *functional* MRI is the specific application of these measurements to detect blood with a high level of oxygen flowing in the brains of experimental subjects performing clearly circumscribed mental tasks. This so-called BOLD (*blood oxygen level dependent*) value is supposed to correlate reliably with circumscribed areas of heightened neural activity, allowing scientists to identify structures in the brain that presumably are involved in bringing about specific mental activities, such as perception, decision-making, memory or emotion (Logothetis, 2008). Gigantic sets of abstract numerical data are generated in this way and, in order to make these large chunks of data manageable, the field developed conventions to plot these data into imagistic depictions of the brain, using different colors to indicate the degree of activation in specific brain regions. The purpose of these depictions is almost exclusively illustrative, as experts usually read out relevant information from more abstract representations of the data (Roskies, 2007).

But it is exactly these colorful brain *images* that account for much of the public as well as professional fascination with the SCAN disciplines. The impression that we are here seeing the brain 'at work' producing mental states seems extremely hard to shake in the face of these depictions, sometimes even despite knowledge to the contrary. Not only can laypeople and poorly informed journalists be swayed into believing that this just *must* be a tremendous advancement of science, professionals from other fields, decision-makers in policy and funding institutions, and certainly students and novices in the brain sciences are often quite

enthralled by the seemingly unshakable impression that fMRI depictions carry tremendous significance (McCabe and Castel, 2008; Racine et al., 2005).

Laypeople coming into contact with the neuroimaging technology are often extremely fascinated and captivated by it. Recent research by medical anthropologists indicates that the machine and its colorful products exert a profound influence, especially on patients and their relatives (Dumit, 2004; Joyce, 2008). Even when subjects are made aware that their brain scans have no medical purpose, the procedure is still viewed as immensely significant and impactful (Cohn, 2012). Apparently, the whole layout of the scanner facilities, the rooms crammed with high-tech equipment, the special security measures (no metal inside the scanner room!), apparently highly competent scientists in white coats operating the machines – all this creates an impression of serious technological advance. The patient's experiences are thus framed by a kind of placebo space, as they are drawn in by the atmospheric qualities of the machine and its setting (Slaby, 2010). This is an effect that might be generalized to much of the current SCAN disciplines and much of the public perception they create.

One tangible effect is that the amount of skepticism directed at alleged findings from neuroimaging, along with the standards of evidence, seem, or at least seemed for some time, lower than in other branches of sciences (Vul et al., 2009; Weisberg et al., 2008). A sign of progress is that the field has recently intensified its methodological self-assessments, headlined by a number of sobering studies of the limitations of the fMRI technology and its established research protocols (see e.g. Carp, 2012; Gonzalez-Castillo et al., 2012; Miller et al., 2012; for a recent overview see Stelzer at al., 2014). But as long as these rather technical discussions do not find a wider audience, the fMRI scanner will continue to function as an *objectivity machine*, seamlessly turning what are mere assumptions and contentious hypotheses into allegedly objective findings, confirmed beyond doubt by what is still widely seen as one of our best scientific methods available. Recognizing the outstanding and ever growing public impact of their field, scientists may be tempted to go with the flow and to seize the moment to acquire more funding, further expand facilities and inscribe their agendas robustly into their institution's long-term research plans – often despite better knowledge with regard to limitations of the technologies and experimental designs (for a scathing recent critique, see Satel and Lilienfield, 2013).

In all sorts of ways, the imaging technology and its publicly visible products shape the way we think of the mind and the brain, the course of research and the design and orientation of research facilities. The instituted practices that surround the fMRI tacitly define what questions are important (because they can be answered by this technology) and what kind of answers will be valued. This is a good example of a successful

cognitive institution exerting power, as some of its institutional features start to shape (and sometimes negatively interfere with) the central scientific aim, viz., in this case, the generation of knowledge about neuronal functioning in humans.

Reformatting Subjectivity – Neuroscience's Socio-cultural Ramifications

The fMRI technology is a case of special interest because it displays central features of the material conditions of science in pronounced ways. On the one hand, it has tremendous advantages as it potentially extends humans' cognitive reach into domains otherwise inaccessible. The shocking complexity of neuronal functioning begins to become available for controlled observation *in vivo*. Likewise, the technique manages to perform a tremendous reduction of complexity by allowing averaging across large chunks of data and by the established practice of plotting individual brain datasets into standardized brain atlases, and, not least, by depicting found activation patterns in the form of vivid images. The MRI scanner is thus a textbook example of what Latour has called 'blackboxing':[4] previously hard-won knowledge is offloaded into a machine whose exact mode of functioning is no longer questioned but simply presupposed by the majority of scientists working with it. This is a prime example of the way science comes to function as a cognitive institution – the cognitive content and cognitive capacities relevant to scientific practice are to a large extent externalized in the technology, and collectivized into communities of specialized experts, exhibiting scientific division of labor – another central principle of the cognitive institution of science.

But then, as much as all this is a great advancement of our collective cognitive capacities, there are downsides as well. Features of the material drivers of scientific practice can run away with their users and take on a life of their own. Here we can focus only on select aspects. Taking the widespread public fascination with neuroimaging as our point of departure, we will discuss specific effects of SCAN on our culture's understanding of subjectivity and selfhood. This is, at the same time, a good example of the way that science as a cognitive institution intersects with other cultural institutions, such as the medical system, the educational system, the legal system, the economy, media, or even the corporate workplace, plus a range of more informal institutions pertaining, for example, to emotions or to construals of selfhood.

A big reason for the widespread fascination with the SCAN disciplines is their specific target phenomena: higher cognitive, emotional and interactive capacities of the human mind. The idea of offering a profoundly scientific approach to the prime levels of human mental functioning promotes the prospect of a bridge between the domain of human experience and the domain of manipulable nature as revealed by the natural

sciences. Finally, or so it might seem, human nature is brought *fully* into the ambit of rigorous scientific investigation. The last bastion withheld from the advancement of scientific naturalism is giving way.

Now, as the actual practice of SCAN is dauntingly complex, difficult and also taxingly incremental, it has become routine to some of its practitioners – especially those already in the public spotlight – to storm ahead of experimental practice by adding a layer of anticipatory, future-oriented philosophy. In fact, it is common practice to couch whatever has been achieved in terms of experimental results within significance-yielding narratives drawn from many sources, but usually oriented in a characteristic fashion around the same recurring themes. Thus, human-level neuroscience is increasingly wedded to a specific naturalistic construal of personhood and subjectivity, as a placeholder and line of flight for what is not yet, but surely one day *will be* revealed by neuroscience's experimental practice. The human mind, in all its fascinating intricacy, lends itself willingly to philosophical stylizations, and, likewise, it isn't too hard to find philosophers happy to offer their services to help formulate a new naturalistic neuro-philosophy of mind, self and personhood. Importantly, these philosophical enrichments are not the erratic work of some individual thinkers, but form part of a much broader movement. In fact, proponents of the new neuro-philosophy can count on a well-organized PR and publishing machinery – churning out bestseller after bestseller in the lucrative genre of 'third culture' (popular) science writing.[5]

A key ingredient of neuroscience-inspired philosophy is that subjective experience, in particular where it pertains to the initiation and control of behavior, is largely illusory, or at best epiphenomenal (Churchland, 2002; Metzinger, 2009). Instead, the actual work in explaining behavior and other phenomena of the personal level is done by vast layers of unconscious neuronal processes – the very processes that presumably begin to yield themselves to rigorous empirical investigation. The mundane self of first-person experience is an illusion, created by post hoc rationalizations of behaviors initiated in the brain, way ahead of the forming of conscious intentions. 'You are nothing but a *pack of neurons*', as Crick (1994) famously put it. Corollary claims follow suit, such as the contention that there is no freedom of will (Wegner, 2002), or that common-sense reality is a neuronal construct, while what is objectively 'out there' is radically different from what is manifest in experience. (On Thomas Metzinger's particularly captivating account, the *real world* 'is just like your physics teacher in high school told you: Out there, in front of your eyes, is just an ocean of electromagnetic radiation, a wild and raging mixture of different wavelengths' [Metzinger, 2009: 20].)

Now, this concerted liquidation, or neurofication of the personal level, including the common-sense lifeworld, is surely not just a matter of philosophical argument and specialist debate. There are much broader

tendencies at work here that warrant a rather different assessment. To put it quite bluntly, regardless of the truth of the neuro-philosophical claims, our social world is already well on the way to becoming a *neuronal world* – societal and cultural transformations in various domains are about to turn it into the very world adequate to the neuro-philosophy just sketched. What may seem quite contentious as philosophy might nevertheless be on the way to become real as culture and practice. In the following, we focus more or less exclusively on those aspects of this trend that have direct links to neuroscience and neuroscience-related themes, but it is important to note that this is part of a broader development going on since roughly 1980. Cognitive science, even in its post-cognitivist era (Clark, 1997), enmeshed with the rise of personal computing, the internet and the spread of mobile communication technology (Fuller and Goffey, 2012; Turkle, 1984), and many technological and economic restructurings in fields such as medicine, education and the workplace have all contributed to creating socio-technological lifeworlds in which characteristic 'nonclassical' understandings of the mind and of personhood become increasingly credible (see Stadler, 2014).

Take the biomedical domain: medical understanding increasingly molecularizes the human body and neurologizes human behavior as it seeks to find causes, drivers and risk factors on the molecular level of organic functioning (Rose, 2007; Rose and Abi-Rached, 2013). Psychiatry increasingly heeds the biological imperative of localizing pathology on the neuronal level, offering pharmaceuticals as the therapy of choice, superseding classical forms of talking therapies that work at the personal level of the patients' life histories and self-understandings (Insel and Quirion, 2005; see Kirmayer and Gold, 2012 for a critique). While actual advances are often slim, discursive practice and aggressive advertising function as placeholders for the soon-to-be-had perfection of molecular medicine. This is nicely illustrated by the way the psychiatric sector manages to raise certain signature illnesses and their neuro-molecularized understandings to public awareness. A good example is depression, which has enjoyed a remarkable media success as a 'chemical imbalance' in the brain, often under cleverly devised labels such as 'burn-out-syndrome', to render it socially acceptable (Ehrenberg, 2010; Healy, 2004). Likewise, ADHD has worked itself to the top of the target list in screening for childhood abnormalities in school settings across the western hemisphere, not least thanks to awareness campaigns promoted by the pharmaceutical industry or by policies devised to implement rigorous biological practice in the mental health sector (Singh, 2011).

Another crucial domain is media (including advertising): modern media-intensive environments increasingly colonize pre-conscious processes such as attention, subliminal perception, instant affective appraisals, quick subconscious decision-making and the like.

Television, the internet, video and computer games, the omnipresence of smart phones and portable networked computers intensify processes at the pre-conscious, automatic, affective levels. Much to the delight of Deleuzians and postmodern affect theorists, the reflectively conscious subject is often circumvented here (see Connolly, 2001; Massumi, 2002; Protevi, 2009). Consider, for example, reports on the use of neuroscientific measures to evaluate the impact of television advertisements (Masters, 2008), or the idea of 'neuro-compression' to target consumers (Friedman, 2012). Thus, present-day media practices signal an obvious line of convergence between current experimental practice in the human sciences and the techno-practical layout of modern lifeworlds. Today's media practices already count on and reckon with the very neuronal 'impulse subject' whose existence is still debated in philosophy (see Fuller and Goffey, 2012).

Likewise, the workplace: while itself increasingly equipped with high-tech media and communication devices and thus rife with the quick reaction time practices, multi-tasking and other routines just described, the modern corporate workplace puts a premium on intensified communication and social skills such as abilities for technology-mediated teamwork and for instant assessment and navigation of social situations, often across global distances. Also here, the reflective, rational, slow and tradition-bound subject is sidelined, giving way to a flexible communicative multi-tasker able to adjust on the fly to new situations, new collaborators, new tasks and surroundings (Boltanski and Chiapello, 2006; Sennett, 2006). Notably, the particular construal of *self* currently championed by social neuroscience – with a focus on automatic mirroring, low-level empathy and mind-reading – neatly corresponds with the ideal skill profile of today's corporate employee (see Hartmann, 2012; Malabou, 2008).

All the developments just sketched bring about new ways of construing oneself 'objectively', in line with the latest findings and construals of the mind offered by neuroscience (in concert with other cognitive science disciplines, evolutionary biology, various branches of psychology and so on). Neuroscience is teaching us how to think differently about the human person, on the supposition that the latter can be reduced to the lowest common mechanism. In this respect, one thinks of everything that is important about being human, including cognitive and emotional life, as in some way centered in and controlled by the brain (see Churchland, 2013). Understand neural processes and you will understand the human. It makes sense, then, to put human subjects into experimental situations and measure the neural events, making these measurements as precise as possible at the most basic level and in the millisecond timescale. On that scale notions of self, intention, agency and free will appear to be more or less illusory products. Such conclusions initially spark debates among philosophers and legal scholars, then enter into the public realm in the

form of public debates or simply summaries of ongoing debates about how such conclusions are already influencing the legal system charged with very practical cognitive tasks of judging and sentencing. In media reports and blogs, the skimpiest details of a complex debate are presented. Neuroscientists are quoted as saying that free will is an illusion; legal scholars are quoted as defending the notion of responsibility; and this, in turn, is questioned by experts in other fields. The scientific data (that a set of neurons activate in a timeframe of 300 milliseconds under certain experimental conditions [see e.g. Libet et al., 1983]) and particular interpretations of that data (that human free will is an illusion) are given equal billing in scientific papers, while in the media and in court, only the interpretations matter and little attention is paid to the debatable suppositions that shape the interpretations (for a critical perspective on the Libet experiments see Schurger et al., 2012).

Accordingly, it is not out of the question that people indeed begin to see themselves as the very neuronal machines, the very molecular automatons apparently championed by the cutting-edge neurocognitive sciences. Narrative self-objectification begins to run ahead of actual research findings, anticipating more than reflecting scientific results (see e.g. Cohn, 2012; Weisberg et al., 2008). There even seems to be a particular willingness, a specific kind of satisfaction gained in construing oneself as powerless, a-rational, subject to uncontrollable external or internal forces that ultimately determine one's fate. Increasingly, we witness what Joseph Dumit calls *objective-self fashioning*:

> The objective-self consists of our taken-for-granted notions, theories, and tendencies regarding human bodies, brains, and kinds considered as objective, referential, extrinsic, and objects of science and medicine…. Furthermore, objective-selfs [sic] are not finished but incomplete and in process. With received-facts, we fashion and refashion our objective-selves. (Dumit, 2004: 7)

This tendency obviously profits from the fascination and appeal exerted by the fMRI scanners and by neuroimaging facilities described above. The 'magical' appeal of the objectivity machine is a central aspect of the dynamics here at work, as it lends the new science a tangible, almost physically graspable credibility and anchors futuristic expectations. In its physical presence, bordering on the hyper-real, the imaging technology acts directly on belief, lending initially vague and little understood ideas the force of conviction. Popular media helps to spread the breathless assumption that we will soon be witnessing an outright neuro-revolution in various areas of society (Lynch, 2009; cf. Heinemann and Heinemann, 2010). Aspects of the objective-self favored by the neurosciences can thus indeed be found *objectively* – they have a robust media presence and

figure as popular shop talk themes before they become resources for individual self-fashioning.

Institutional Entanglements

In terms of our proposal to understand neuroscience as a cognitive institution, the descriptions so far given have led to a significant increase in complexity. Instead of just one, clearly circumscribed institution with clear-cut rules, procedures, practices and participants, we have found neuroscience to be the point of intersection of *various* different cultural and social and cognitive institutions and thus of the activities, skills and interests of a whole array of agents, representatives and stakeholders.

Obviously, it would be a mistake to assume a sharp boundary between the alleged 'inner' domain of neuroscience – for example, represented by the doings of professional neuroscientists, the equipment found in neuroscience labs, experimental procedures, conferences, publication and peer-review practices, etc. – and what seems merely adjacent to it, for example media, various domains of application (e.g. medicine or law or education), or the conceptual commentaries provided by (neuro-)philosophy. A closer look at the day-to-day work of neuroscientists reveals the depth of the entanglement, the intensity of interaction between this presumed 'inside' of neuroscience and the various users, sponsors, mediators, facilitators presumably on its outside. In fact, neuroscientists in leading positions spend tremendous amounts of their work time with activities such as fundraising (applying for grants, contacting and talking to potential funders or sponsors), PR (giving interviews, popular lectures, appearing in talk shows, etc.), institutional networking and research policy (administering institutes, creating strategic alliances, engaging in institutional politics, etc.), all very much encouraged by the universities and institutions to which they belong. Moreover, an activity such as academic teaching – a key part of what most scientists regularly do – often reaches broader audiences and encompasses broader ways of discursive relatedness than just what strictly belongs to an alleged core of a scientific discipline. But this is not all. As we will see in the remainder of this section, even activities that might clearly belong to 'research proper' are, to an important extent, infiltrated by various elements and tendencies from other institutional practices, rendering hopeless any attempt to separate what is actual *scientific* work from what are merely additional extras. Accordingly, we hold that the cognitive institutions of science are thoroughly and constitutively interrelated with many other social institutions, including other cognitive institutions.

There are many different ways in which these institutional entanglements become manifest in scientific practice. Start with the selection of research themes. Why all this work on social cognition and intersubjectivity (after the long dominance of methodological individualist

approaches)? Why the widespread turn to emotion and affect in the 1990s and onward? Why the focus on some pathologies and the neglect of others? Why are we witnessing entire theme industries, where, instead of researchers globally working on vastly different issues relevant to human affairs, we see a 'tight clumping' of research activities around a small number of topics that are often worked to death (again: emotion, empathy, attention, certain mental illnesses, etc.)? One reason might be the involvement of powerful stakeholders (such as Big Pharma or the ICT sector) behind some of the themes mentioned – for example, the recently inaugurated, massively funded Human Brain Project of the EU would be unthinkable without massive 'digital agenda' involvement (see Markram et al., 2011). Relatedly, the influence of a relatively small number of funding agencies impacts research agendas by pre-assigning funds to a few select so-called 'key' issues. This then points to a more indirect force behind such institutional policies: the capitalist logic of cycles of boom and decline, which can be hard to resist for anyone who wants to become a serious player in a given field. In this way, self-referential circles are set up, hyping up specific themes while neglecting others that would equally warrant intensive research.

This is not the only way in which structuring effects from the capitalist economy make themselves felt within neuroscience. Besides the involvement and influence from corporate players that may function as direct sponsors and customers of neuroscience, several more indirect effects can be observed. The most general and most important is the fact that SCAN increasingly bears the marks of a so-called 'economy of hope' (Novas, 2006). As much of today's economy is no longer an economy of production but an economy of speculation – that is, consisting of bets on economic or monetary developments in the future – much the same seems to hold for the economic uptake of neuroscience. Instead of neuroscientific applications *already available*, much of the economic appeal of the field lies in its *promissory character* (Hagner and Borck, 2001). The standard tune goes like this: not quite yet, but *very soon* we will have applications with world-changing potential, promising astronomical revenues – be it reliably working fMRI-based lie detectors, effective neuro-implants and prostheses, next-generation psycho-pharmaceuticals that will help cure mental illnesses, or brain–computer interfaces set to revolutionize rehabilitation, communication or gaming. Thus, human-level neuroscience is very much in line with the speculative economy in offering bets on the future: scenarios just appealing and credible enough to persuade investors to open their cheque books. To be sure, as of yet, only some neuroscientists directly engage investment capital (often in the form of corporate spin-offs). But the formatting effects of this trend – patterned after the bio-capital boom surrounding genetics in the 1990s and early 2000s (Cooper, 2008; Sunder Rajan, 2006) – reach much farther. For example, aspects of the future-directed market logic have

an effect on the selection of research themes (maximize impact and attention!), the mode of presentation of results (claim the most and promise even more!), and the mode of interaction between scientists and the media (frequent and intensive!). Likewise, the forms of presentation that favor the spectacular and futuristic promises have reformatted the character of scientific presentations and teaching. The trend is towards *edutainment* (see Heinemann, 2012).

But the structuring effects of the capitalist logic reach further than into the presentation styles of some researchers. Governing boards of research institutions and universities, as well as policy-makers, likewise believe in the bright outlook of the neurosciences. Accordingly, efforts are undertaken to strengthen the institutional status of the SCAN disciplines, for example by setting up new research facilities or enlarging existing ones. As these measures usually include the purchase and installation of expensive hardware such as MRI scanners and the special facilities required by them, these processes are set to leave lasting marks upon the research landscape. Notably, this neuro-expansion trend has effects upon other disciplines, often signaled by the advent of the new neuro-fields in social sciences and humanities, for example, *neuro-aesthetics, neuro-politics, neuro-literary criticism* or *neuro-theology* (Ortega and Vidal, 2010; Pickersgill and Van Keulen, 2012). This disciplinary reformatting often implies profound changes to the research agendas, practices and standards of the fields in question. For example, aesthetic disciplines now increasingly focus on the psychological mechanisms allegedly involved in art reception, instead of assessing the aesthetic qualities and contents of artworks according to established disciplinary and genre-specific criteria (Vidal, 2009). Fields belonging to the humanities are first psychologized as the focus is shifted towards individual psychological mechanisms and, second, neuralized in that an effortful search for the 'neural correlates' of these mechanisms is initiated (*neuro-politics* is another good example for this trend – see the essays in Vander Valk, 2012, and especially Slaby et al., 2012, for a critique).

To be clear, we are not denying that some of these reforms in academic orientations might open up interesting avenues for research. Nor do we deny that traditional and previously established disciplines and ideas have been pushed and pulled into shape by earlier technologies and scientific discoveries. We do want to suggest, however, that taking a critical perspective on these changes, and the possible dangers, distortions, and narrowing effects lurking in them, is important. The dominance of neuropsychological paradigms might easily shut off other viable ways of inquiry into human reality, especially those that are initially less flashy, less publicly appealing and less promising in terms of corporate revenues.

In sum, we have seen that the bundle of social and cognitive institutions that jointly make up SCAN is inseparably entangled with many other social and cultural institutions and institutional structures,

strategies and trends. We have likewise seen that these various institutional practices, operating procedures, rules and standards have profound effects on the way neuroscience is practiced, communicated, taught and spread to other fields. While surely not all (and not all equally) negative, these trends and developments call for critical scrutiny. It is important to develop a perspective from which practitioners, scholars and educated laypeople are able to analyze potentially distorting tendencies and measure their impacts upon the institutional structure of neuroscience, not least assessing them with regard to what – we hope – is still neuroscience's prime scientific objective: to understand and explain neuronal functioning in humans. Accordingly, we close our article by discussing briefly the need for critique, both generally with regard to cognitive and other institutions and specifically with regard to neuroscience.[6]

Outlook: The Indispensability of Critique

Neuroscience – like other scientific disciplines – is a powerful collective tool (or rather, set of tools) capable of extending our understanding of human nature in helpful ways. Its instruments and materials as well as its collaborative social organization as a distributed practice incorporate large bodies of knowledge, both procedural and propositional, plus many sorts of practical skills and crafts. Various institutional conventions (e.g. those pertaining to data processing and standardization, peer review, conferences, etc.) facilitate the scientists' work and help them coordinate their many, often widely distributed activities. Crucially, as such, the social and cognitive institutions that help to constitute the SCAN knowledge are densely enmeshed with a broad range of other institutions and their complex modes of operation, their goals, rules and styles. With regard to these institutional entanglements, we have provisionally charted several different but interrelated developments, in each case already with a view towards potentially distorting or otherwise problematic tendencies. Thus, the foregoing already was in part an exercise in critical neuroscience.

It is important to see this as an inevitable dialectic. The institutional processes and interrelations here charted belong to the indispensable enabling conditions of many important human achievements. It is a key mark of the human to be able to set up and maintain such powerful systems of entangled institutions that help individuals achieve more, cognitively as well as practically. However, with all those complex social and cognitive practices – such as the intersecting of various spheres of institutional practice with their distinct functioning principles, such as the economy, media, the law or the medical sector – inevitably comes the likelihood of distortions, overly one-sided developments, abuses of power, undue influences. These problematic

tendencies and structural pathologies appear in many forms, often, for example, as an over-extension of the principles and orientations of one institution taking over other institutions. This is the case, obviously, when capitalist profit orientation and futuristic speculation 'invades' other domains such as science, overriding some of science's own goals and modes of functioning (Mirowski, 2011). Likewise, when requirements of the powerful media sector – such as catchy presentation styles focusing on sensationalistic themes – influence the theme selection, communication and presentation styles employed in scientific cognition. Science itself can likewise unduly influence other institutional spheres, for example when a discipline such as SCAN prematurely exports its alleged 'results' – often in the form of contentious worldviews standing in for scientific results not yet attained – to other areas of society, such as the medical sector, education or social policy. Neuroscience then risks abusing the trust that the public places in it, satisfying its hunger for public attention, acclaim, institutional status and economic profit – orientations that themselves reflect influences from other institutions – more than its specifically scientific objectives. This can have wide-ranging consequences, for example when a narrow physiological understanding of mental illness begins to shut down other useful ways to come to terms with psychopathology, or when contentious educational or social policy practices are implemented despite the lack of robust evidence in their favor.

What we propose is not to even think that we could or should draw clear boundaries or set up barriers between these various institutions. Indeed, even if simplistic media distortions concerning neuroscientific results can be damaging to self-understanding, the cure is not to keep the media away from neuroscience. Rather, the proposal is to develop a toolkit of critical practices, helping scientific practitioners and commentators chart problematic developments, assess their actual and potential impacts, and work out alternatives. Where possible, this would elevate public debate to a critical level. In this sense, critique is a kind of cognitive maintenance routine for complex institutions, required to provide reflexive insights into the many intersecting institutional structures that make up modern lifeworlds and that scaffold individual cognition and behavior. On the most global plane, critique is the form of self-reflexivity needed to guard against the tendency of institutions to take on lives of their own, running away from the interests and intentions of the people in whose service they have been set up. The critical theorist Axel Honneth uses a helpful phrase, holding that critique aspires to detect, describe and work towards alleviating *pathologies of reason*, where 'pathology' is meant to refer to a structural deficit in institutional organization, causing suffering or preventing individuals from leading rich and satisfied lives (Honneth, 2009). Importantly, this critical activity has to include the constantly renewed reflection on the general

significance and specific purposes of institutional arrangements, and also a recurrent scrutiny of the normative grounds – which are subject to historical shifts – on which both current institutional set-ups *and* attempts to critically assess them are based (Hartmann and Honneth, 2006). As we hope to have made clear, the focus on institutions, and notably on cognitive institutions, is a key move as it helps us to gain traction on the most effective among the processes that shape, mold and police contemporary human life forms. However, this is not the place to describe in detail the various forms that such a self-reflexive practice of critique might take in the case of neuroscience (we developed some of this elsewhere; see Gallagher, 2012, and the other contributions in Choudhury and Slaby, 2012).

Instead, we close by discussing a more fundamental issue. Let us return to the current reconception or reformatting of subjectivity and selfhood in Western societies that is in part enabled by the recent successes of the neurosciences. The human subject as pictured by neuroscience is ruled by subpersonal processes, programmed by evolution to respond reflex-like to environmental affordances, blindly executing a large number of perceptual, affective and behavioral routines. What if it was indeed true that the rational and responsible subject championed by the Enlightenment turns out to be an illusion? What would *this* entail for the possibility of reflexivity and critique as just outlined? This, in effect, points to the possibility of a thorough counter-critique courtesy of neuroscience and neuro-philosophy, preventing a critical neuroscience from getting off the ground. What if there simply was no human subject, rational, reflective and self-consciously autonomous enough to initiate and carry out the kind of critical scrutiny demanded by the critical neuroscience approach? What if the very conditions required for critical reflexivity of scientific institutions were simply not fulfilled, and this was already shown by a reductionistic neuroscience that conceives of the mind as entirely brain-bound? Does this mean that a critical neuroscience is only possible on condition of the *falsity* of the anti-rationalistic, anti-reflective conception of the human subject championed by neuroscience and neuro-philosophy?

Now, we do indeed believe that this construal of the human subject will turn out to be empirically false (or, in any case, unduly simplistic and one-sided). But we do not need this to be established prior to our call for a critical self-reflective perspective. It is important to note that the task of a critical neuroscience is not to somehow show that the conception of the human subject put forward by SCAN is false. Indeed, even if we assumed it to be true, this would not mean that a critical neuroscience is impossible. This returns us to the beginning of our article. The whole point of the socially extended mind hypothesis can be put like this: almost regardless of what the 'naked' and isolated human individual, left to his on-board, brain-bound devices, is capable of – what counts in

determining the cognitive capacities of the human subject is the intersubjective cognitive systems that result when the embodied individual is coupled to various cognition-enabling environmental structures. On the perspective taken here, it is not isolated brains, but fully embodied persons coupled to cognitive institutions (which include, most fundamentally, language and communicative practices with others) that constitute humans' full cognitive, rational capacities. To conceive of the mind as socially extended is to conceive of it as something that is irreducible to neuronal processes and, accordingly, in a way that critically challenges overly reductionistic tendencies of SCAN disciplines. Thus our answer to the neuro-philosophical challenge: the kind of rational self-reflective, accountable practice required to carry out the critical work we have proposed is itself in part an institutional reality, enabled and sustained by the collective resources and capacities provided by collective institutional practices. The rational human subject is not an exclusively biological entity – it is an entity coupled to other biological individuals and various cognition-enabling institutions, tools, procedures and practices. This very institutional reality endows humans with the capacities needed to distance themselves from and critically question ongoing practice and engagement – capacities that animals, lacking these collective resources, are utterly unable to develop. The moral of this is, of course, a critical message addressed directly to neuroscientists: Don't take the individual biological system that you study experimentally for the full human person! Either find ways to encompass the workings of cognitive institutions, mind extensions and social-normative practices in your research, or stop declaring yourself able to study the human mind in its full scope. The mind is wider than the brain (Rose, 2012).

Notes

1. The parity principle states: 'If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process' (Clark and Chalmers, 1998: 8).
2. A more technical definition of dynamic systems coupling is provided by Di Paolo and De Jaegher (2012):

> two systems are said to be coupled when parametrical and other structural descriptions of the laws of transformation of states in one of them have a functional dependence on the state variables of the other, which may be non-linear, piece-wise, state-dependent, and time-varying (in which case we call the coupling 'dynamical'). Coupling can be unidirectional or mutual.

3. We do not have the space to present a full exposition of the approach to critique that we champion as a key ingredient of critical neuroscience.

Suffice it to say that we roughly follow the broader socio-political mission of Frankfurt School critical theory, and also their particular approach of critique as rational self-reflection on institutional practices in democratic societies. Thus, as we suggest here, a key task is to reconstruct, analyze and assess the normative foundations of institutions operative in society. In particular, it is crucial to assess whether the actual functioning of the institutions in question is living up to the normative orientation legitimizing their operations. We can only provide a rough sketch of such an encompassing assessment of neuroscience as a cognitive institution within the limited space of this paper. On the understanding of 'critique' that we draw on, see Geuss (1981), Honneth (2009) and, for an assessment of the impacts of recent transformations in western capitalism on the normative foundations of critique, Hartmann and Honneth (2006). A more thorough discussion of the notion of 'critique' relevant to critical neuroscience is offered in Choudhury and Slaby (2012: 1).

4. 'The way scientific and technical work is made invisible by its own success. When a machine runs efficiently, when a matter of fact is settled, one need focus only on its inputs and outputs and not on its internal complexity. Thus, paradoxically, the more science and technology succeed, the more opaque and obscure they become' (Latour, 1999: 304).
5. We refer here to authors such as Patricia Churchland, Richard Dawkins, Daniel Dennett, Steven Pinker, Craig Venter, Daniel Kahneman, Thomas Metzinger, Marvin Minsky and many others, plus several popular neuroscientists – such as Antonio Damasio, Michael Gazzaniga, Vilayanur Ramachandran and Joseph LeDoux – whose writings draw much of their appeal from substantive philosophical claims on top of spectacular tales from the neurological clinic or the neuroscience lab bench. It is noteworthy that most of these authors are under contract with literary agent John Brockman, founder of the Edge Foundation (edge.org), who coined the label 'third culture' to subsume his illustrious clientele (Brockman, 1995). See Stadler (2014) for a pointed historical analysis and critical discussion of the Brockmanites, and also Žižek (2002) for a discussion of the 'third culture' movement.
6. Various articles in Choudhury and Slaby (2012) and in a forthcoming Special Topic (on critical neuroscience) in *Frontiers in Human Neuroscience* provide examples of the kind of critical work we envisage here. Comparable approaches are found in several contributions to Ortega and Vidal (2010) and in Littlefield and Johnson (2012).

References

Adams F and Aizawa K (2008) *The Bounds of Cognition*. Malden, MA: Blackwell.

Adams F and Aizawa K (2010) Defending the bounds of cognition. In: Menary R (ed.) *The Extended Mind*. Cambridge, MA: MIT Press, pp. 67–80.

Bassolino M, Serino A, Ubaldi S and Ladavas E (2009) Everyday use of the computer mouse extends peripersonal space representation. *Neuropsychologia* 48(3): 803–811.

Boltanski L and Chiapello E (2007) *The New Spirit of Capitalism*, trans. Elliot G. London: Verso.

Brockman J (1995) *Third Culture: Beyond the Scientific Revolution*. New York: Simon and Schuster.
Carp J (2012) On the plurality of (methodological) worlds: Estimating the analytic flexibility of 627 fMRI experiments. *Frontiers in Neuroscience* 6. doi:10.3389/fnins.2012.00149.
Choudhury S and Slaby J (eds) (2012) *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*. Chichester: Wiley-Blackwell.
Churchland PS (2002) *Brain-wise: Studies in Neurophilosophy*. Cambridge, MA: MIT Press.
Churchland PS (2013) *Touching a Nerve: The Self as Brain*. New York: W.W. Norton.
Clark A (1997) *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.
Clark A (2008) *Supersizing the Mind: Reflections on Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.
Clark A (2009) Spreading the joy? Why the machinery of consciousness is (probably) still in the head. *Mind* 118: 963–993.
Clark A and Chalmers D (1998) The extended mind. *Analysis* 58(1): 7–19.
Cohn S (2012) Disrupting images: Neuroscientific representationsin the lives of psychiatric patients. In: Choudhury S and Slaby J (eds) *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*. Chichester: Wiley-Blackwell, pp. 179–193.
Connolly WE (2001) *Neuropolitics: Thinking, Culture, Speed*. Minneapolis, MN: University of Minnesota Press.
Cooper M (2008) *Life as Surplus: Biotechnology and Capitalism in the Neoliberal Era*. Seattle: University of Washington Press.
Cooter R and C Stein (2013) *Writing History in the Age of Biomedicine*. New Haven, CT: Yale University Press.
Crane T (1998) Intentionality as the mark of the mental. In: O'Hear A (ed.) *Contemporary Issues in the Philosophy of Mind*. Royal Institute of Philosophy Supplement (No. 43). Cambridge: Cambridge University Press.
Crick F (1994) *The Astonishing Hypothesis: The Scientific Search for the Soul*. New York: Scribner.
De Jaegher H, Di Paolo E and Gallagher S (2010) Does social interaction constitute social cognition? *Trends in Cognitive Sciences* 14(10): 441–447.
Dewey J (1916) *Essays in Experimental Logic*. Chicago: University of Chicago Press.
Di Paolo E (2009) Extended life. *Topoi* 28: 9–21.
Di Paolo E and De Jaegher H (2012) The interactive brain hypothesis. *Frontiers in Human Neuroscience* 6. doi: 10.3389/fnhum.2012.00163.
Dumit J (2004) *Picturing Personhood: Brain Scans and Biomedical Identity*. Princeton, NJ: Princeton University Press.
Ehrenberg A (2010) *The Weariness of the Self. Diagnosing the History of Depression in the Contemporary Age*. Montreal: McGill–Queen's University Press.
Friedman W (2012) CBS: TV spots more effective using neuroscience techniques. *Media Daily News*, 28 June. Available at: http://www.mediapost.com/

publications/article/177673/cbs-tv-spots-more-effective-using-neuroscience-te.html (accessed 30 September 2014).
Fuller M and Goffey A (2012) *Evil Media*. Cambridge, MA: MIT Press.
Gallagher S (2012) Scanning the lifeworld: Towards a critical neuroscience of action and interaction. In: Choudhury S and Slaby J (eds) *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*. Chichester: Wiley-Blackwell, pp. 85–110.
Gallagher S (2013) The socially extended mind. *Cognitive Systems Research* 25–26: 4–12.
Gallagher S and Crisafi A (2009) Mental institutions. *Topoi* 28(1): 45–51.
Gallagher S and Miyahara K (2012) Neo-pragmatism and enactive intentionality. In: Schulkin J (ed.) *Action, Perception and the Brain*. Basingstoke: Palgrave Macmillan, pp. 117–146.
Geuss R (1981) *The Idea of a Critical Theory: Habermas and the Frankfurt School*. Cambridge: Cambridge University Press.
Gonzalez-Castillo J, Saad JZ, Handwerker DA, Inati SJ, Brenowitz N and Bandettini PA (2012) Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis. *Proceedings of the National Academy of Sciences* 109(14): 5487–5492.
Hagner M and Borck C (2001) Mindful practices: On the neurosciences in the twentieth century. *Science in Context* 14(4): 507–510.
Hartmann M (2012) Against first nature: Critical theory and neuroscience. In: Choudhury S and Slaby J (eds) *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*. Chichester: Wiley-Blackwell, pp. 67–84.
Hartmann M and Honneth A (2006) Paradoxes of capitalism. *Constellations* 13(1): 41–58.
Healy D (2004) *Let Them Eat Prozac: The Unhealthy Relationship between the Pharmaceutical Industry and Depression*. New York: New York University Press.
Heinemann LV and Heinemann T (2010) Optimize your brain! Popular science and its social implications. *BioSocieties* 5: 291–294.
Heinemann T (2012) *Populäre Wissenschaft. Hirnforschung zwischen Labor und Talkshow*. Göttingen: Wallstein.
Honneth A (2009) *Pathologies of Reason: On the Legacy of Critical Theory*, trans. Ingram J. New York: Columbia University Press.
Hutto D and Myin E (2013) *Radicalizing Enactivism*. Cambridge, MA: MIT Press.
Insel TR and Quirion R (2005) Psychiatry as a clinical neuroscience discipline. *Journal of the American Medical Association* 294: 2221–2224.
Joyce KA (2008) *Magnetic Appeal: MRI and the Myth of Transparency*. Ithaca, NY: Cornell University Press.
Kirmayer LJ and Gold I (2012) Re-socialising psychiatry: Critical neuroscience and the limits of reductionism. In: Choudhury S and Slaby J (eds) *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*. Chichester: Wiley-Blackwell, pp. 307–330.
Latour B (1999) *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.
Latour B (2004) Why critique has run out of steam: From matters of fact to matters of concern. *Critical Inquiry* 30: 225–248.

Libet B, Gleason CA, Wright EW and Pearl DK (1983) Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain* 106(3): 623–642.
Littlefield M and Johnson J (2012) *The Neuroscientific Turn: Transdisciplinarity in the Age of the Brain*. Ann Arbor, MI: University of Michigan Press.
Logothetis NK (2008) What we can do and cannot do with fMRI. *Nature* 453(7197): 869–878.
Lynch Z (2009) *The Neuro Revolution: How Brain Science is Changing Our World*. New York: St Martin's Press.
McCabe DP and Castel AD (2008) Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition* 107: 343–352.
Maguire EA, Gadian DG, Johnsrude IS, Good CD, Ashburner J, Frackowiak RS, et al. (2000) Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences* 97(8): 4398–4403.
Malabou C (2008) *What Should We Do with Our Brain?* New York: Fordham University Press.
Markram H, Meier K, Lippert T, Grillner S, Frackowiak R, Dehaene S, et al. (2011) Introducing the Human Brain Project. *Procedia Computer Science* 7: 39–42.
Massumi B (2002) *Parables for the Virtual: Movement, Affect, Sensation*. Durham, NC: Duke University Press.
Masters K (2008) Neuroscience helps marketers judge ads' impact. National Public Radio, 22 July. Available at: http://www.npr.org/templates/story/story.php?storyId=92768079 (accessed 30 September 2014).
Menary R (2007) *Cognitive Integration: Mind and Cognition Unbounded*. Basingstoke: Palgrave Macmillan.
Menary R (2009) Intentionality, cognitive integration, and the continuity thesis. *Topoi* 28: 31–43.
Metzinger T (2009) *The Ego-Tunnel: The Science of the Mind and the Myth of the Self*. New York: Basic Books.
Miller MB, Donovan C-L, Bennett CM, Aminoff EM and Mayer RE (2012) Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals. *NeuroImage* 59: 83–93.
Mirowski P (2011) *Science-Mart: Privatizing American Science*. Cambridge, MA: Harvard University Press.
Novas C (2006) The political economy of hope: Patients' organizations, science and biovalue. *BioSocieties* 1: 289–305.
Ortega F and Vidal F (eds) (2010) *Neurocultures: Glimpses into an Expanding Universe*. Berlin: Peter Lang.
Pickersgill M and Van Keulen I (2012) *Sociological Reflections on the Neurosciences*. Bingley: Emerald.
Pickering A (1995) *The Mangle of Practice: Time, Agency and Science*. Chicago: University of Chicago Press.
Protevi J (2009) *Political Affect: Connecting the Social and the Somatic*. Minneapolis, MN: University of Minnesota Press.
Racine E, Bar-Ilan O and Illes J (2005) fMRI in the public eye. *Nature Reviews Neuroscience* 6: 9–14.

Rheinberger H-J (1997) *Toward a History of Epistemic Things: Synthesizing Proteins in a Test Tube*. Stanford, CA: Stanford University Press.
Rose N and Abi-Rached J (2013) *Neuro: The New Brain Sciences and the Management of the Mind*. Princeton, NJ: Princeton University Press.
Rose N (2007) *The Politics of Life Itself: Biomedicine, Power, and Subjectivity in the Twenty-First Century*. Princeton, NJ: Princeton University Press.
Rose S (2012) The need for a critical neuroscience: From neuroideology to neurotechnology. In: Choudhury S and Slaby J (eds) *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*. Chichester: Wiley-Blackwell, pp. 53–66.
Roskies AL (2007) Are neuroimages like photographs of the brain? *Philosophy of Science* 74: 860–872.
Rouse J (1987) *Knowledge and Power: Toward a Political Philosophy of Science*. Ithaca, NY: Cornell University Press.
Rowlands M (2009) The extended mind. *Zygon* 44(3): 628–641.
Rupert R (2009) *Cognitive Systems and the Extended Mind*. New York: Oxford University Press.
Satel S and Lilienfield SO (2013) *Brainwashed: The Seductive Appeal of Mindless Neuroscience*. New York: Basic Books.
Schurger A, Sitt JD and Dehaene S (2012) An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences* 109(42): E2904–E2913.
Sennett R (2006) *The Culture of the New Capitalism*. New Haven, CT: Yale University Press.
Singh I (2011) A disorder of anger and aggression: Children's perspectives on ADHD in the UK. *Social Science & Medicine* 73(6): 889–896.
Slaby J (2010) Steps towards a critical neuroscience. *Phenomenology and the Cognitive Sciences* 9: 397–416.
Slaby J, Haueis P and Choudhury S (2012) Neuroscience as applied hermeneutics: Towards a critical neuroscience of political theory. In: Vander Valk F (ed.) *Essays on Neuroscience and Political Theory*. New York: Routledge, pp. 50–73.
Stadler M (2014) Neurohistory is bunk? The not-so-deep history of the post-classical mind. *Isis* 105(1): 133–144.
Stelzer J, Lohmann G, Mueller K, Buschmann T and Turner R (2014) Deficient approaches to human neuroimaging. *Frontiers in Human Neuroscience* 8(462).
Sunder Rajan K (2006) *Biocapital: The Constitution of Postgenomic Life*. Durham, NC: Duke University Press.
Thompson E (2007) *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.
Turkle S (1984) *The Second Self: Computers and the Human Spirit*. New York: Simon and Schuster.
Vander Valk F (ed.) (2012) *Essays on Neuroscience and Political Theory*. New York: Routledge.
Varela F, Thompson E and Rosch E (1991) *The Embodied Mind*. Cambridge, MA: MIT Press.
Vidal F (2009) Brainhood, anthropological figure of modernity. *History of the Human Sciences* 22: 5–36.

Vul E, Harris C, Winkielman P and Pashler H (2009) Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science* 4: 274–290.
Weisberg DS, Keil FC, Goodstein J, Rawson E and Gray JR (2008) The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience* 20(3): 470–477.
Wegner DM (2002) *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
Wheeler M (2010) Minds, things and materiality. In: Malafouris L and Renfrew C (eds) *The Cognitive Life of Things: Recasting the Boundaries of the Mind*. Cambridge: McDonald Institute for Archaeological Research, pp. 29–38.
Young A (2012) Empathic cruelty and the origin of the social brain. In: Choudhury S and Slaby J (eds) *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*. Chichester: Wiley-Blackwell, pp. 158–176.
Žižek S (2002) Cultural studies versus the 'third culture'. *South Atlantic Quarterly* 101(1): 19–32.

Jan Slaby is Assistant Professor in Philosophy of Mind at Free University, Berlin. He has published on emotion and affect, agency, embodiment, on issues in the philosophy of psychiatry and on critical neuroscience. He is co-founder of the critical neuroscience initiative (www.critical-neuroscience.org).

Shaun Gallagher is the Lillian and Morrie Moss Professor of Excellence in the Department of Philosophy at the University of Memphis. He has secondary appointments at the University of Hertfordshire and the University of Wollongong. He also holds the Anneliese Maier Research Award from the Humboldt Foundation (2012–17).