PSIWORLD 2013

Personal identity, functionalism and the extended mind

Marius M. Stanciu*

University of Bucharest, Faculty of Psychology and Educational Sciences, 90 Șos Panduri, Bucharest, Romania

Abstract

In the last thirty years, advancements made in neurobiology and computer science have profoundly changed the ways in which we conceive and regard ourselves. Within this context, metaphysical questions about personal identity become intertwined with real, practical concerns and ethical considerations. In this paper we will examine the extended mind thesis (Clark & Chalmers, 1998) and what it entails for personal identity. Adopting a psychological continuity view restated in functionalist terms, we conclude that mind extension is currently impossible, but might be possible in the future. Ultimately, we maintain that functionalism paints a very optimistic picture about our psychological persistence conditions.

© 2014 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
Selection and peer-review under responsibility of Romanian Society of Applied Experimental Psychology.

Keywords: Identity, Extended Mind, Psychological Continuity, Functionalism, Consciousness.

1. Introduction

From ancient times to the present day, philosophy of mind has presented us with some of the most intractable problems imaginable. Leaving aside the “hard” problem of consciousness (Chalmers, 1995), issues regarding mental causation, intentionality and our seemingly effortless access to our private mental states have generated not only an enormous body of literature, but also vigorous debates that continue to the present day. In numerous ways, the topic

* Marius M. Stanciu. Tel.: +40.0744.535.123; fax: +4.031.409.3254.
E-mail address: marius_ms_stanciu@yahoo.com
of the present paper touches all of these subjects, being preoccupied with one of the most intimate, if not the most intimate concept that we can have – namely, our personal identity.

In the last thirty years, advancements made in neurobiology, as well as computer science, have profoundly changed the ways in which we regard and conduct ourselves. These broad lines of research have ultimately given birth to new perspectives over the issue of personal identity, either extending the mind over parts of the environment, or reducing the notion of a person to its biological, animal form. Within the given context, the main purpose of this paper is to investigate the necessary and sufficient conditions that ensure the persistence of a person such as to avoid the problematic consequences of both the extended mind thesis and the somatic approach (e.g. Olson, 1997, for the latter). As such, our concern here will be rather metaphysical than epistemic in character.

While giving an exhaustive treatment of the literature is impossible due to the limited space, we will begin here by following the distinction made by Parfit (1982) between the Simple View of personal identity and the more articulate, Complex View, that takes psychological or somatic continuity as relevant criteria for judging identity relations. In the second section we will briefly discuss the extended mind thesis (Clark & Chalmers, 1998), or what Adams and Ainszawa (2001) called “contingent transcranialism”, further exposing the consequences of this approach in drawing the lines of the self. Finally, in the last section we will argue that even though extending the boundaries of the mind through artificial means is not only a logical possibility, but also a nomological one, such extensions are currently impossible. From a functionalist point of view, extension requires the existence of an underlying single mechanism that causally bridges the mind states occurring on different vehicles. If this is a desirable option is, of course, another question which we cannot give a full and elaborate treatment here.

2. Psychological continuity as necessary and sufficient for persistence

The Simple View / Complex View distinction was introduced by Derek Parfit in the 1980s in order to describe and classify the contrasting perspectives on personal identity that have been advanced in the Western tradition of thought. While this dichotomy certainly has its shortcomings, it is nonetheless a useful intellectual tool that provides important insight into the problem. In broad strokes, following Parfit, according to the Simple View of the non-reductionist tradition, personal identity is a separate “further fact” (Parfit, 1982, p. 227) that cannot be reduced to mental or physical continuity (e.g. Chrisholm, 1976). Often, partisans of this perspective also adopt the idea that we are res cogitans substances, immaterial Cartesian egos or spirits that will continue to exist well after our physical bodies are destroyed, but some may also consider the possibility that we are actually physical entities of a different kind, yet undiscovered by physics (Parfit, 1984, p.210). Strictly speaking, the distinction introduced earlier is actually more concerned with the question of a person’s persistence conditions or what matters for survival than with the metaphysical nature of a person. Some philosophers (e.g. Wilkes, 1998, vii) consider such problems inextricably linked, but here we prefer to be reserved on this issue. However improbable and strange it may seem in the present age dominated by physicalism, the Simple View is “almost certainly the default folk view” (Matthews, 2010, p. 183). Understanding how thinking entities can persist in time but not in space is enormously difficult and, due to many other considerations coming from empirical research, we will leave aside the Simple View and concentrate on the much more articulate Complex perspective.

In nuce, whether we talk about a somatic or a psychological approach, those who propose a Complex understanding of identity relations are “reductionists” in the sense that they consider sub-personal facts to determine necessary and sufficient conditions for the persistence of a person. The most popular such view is undeniably the psychological continuity one, although recently, Animalism has gathered a sufficient amount of supporters. Carrying a certain Lockeian flavour in their analyses, adepts of the psychological approach follow the British empiricist in considering that what matters for personal continuity ultimately revolves around that entity being a “thinking, intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places” (Locke, 1689, II.xxvii.9). Of course, as Reid and Butler have discovered in the eighteenth century (see Perry, 2002, p. 84), appeals to memory cannot get us very far because it is an unreliable mental faculty. First, forgetting an episode from one’s life violates transitivity, which is impossible for identity relations according to the formal rules of logic. Real life conditions such as retrograde and anterograde amnesia further complicate the picture, leading us to believe that an individual either becomes another person after the accident (in the first case) or
becomes an indefinite number of persons only loosely connected to the original one (in the second case). Secondly, the memory continuity principle is trivial in the sense that remembering an experience from your life presupposes implicitly that you are the same person who had that experience. Although this fact does not undermine in any way the memory connection necessity for identity, it doesn’t lead to an informative theory either.

Usually, attempts to respond to the first problem have shifted interest from explicit to implicit memory, while avoiding the circularity of the second led to the introduction of “quasi-memory” (Shoemaker, 1970) – a concept which operates a distinction between the “content” and the “experience” of a memory. While prima facie such a notion may seem unnatural or counterintuitive, it offers not only an elegant solution to the problem, but it is also compatible with current neuropsychological and neuropsychiatric findings (Northoff, 2000). Interesting as they are, we will not pursue here such approaches further due mainly to two reasons. First, they are still considered controversial conceptions in some circles. Secondly, psychological connectedness offers a much broader and plausible picture for the persistence conditions of persons without entering such muddy waters.

Reduced to its core, the psychological continuity view simply says that a person P at time t is identical with a person P’ at t’ if and only if: (1) the mental states of P’ are causally dependent on the mental states of P in an appropriate fashion (i.e. following a specific relation); and (2) at no time in the interval from t to t’ does this series of mental states “branch” (Davis, 1998, pp. 781-782). There are, of course, many ways of thinking about mental states. Assuming a type-identity framework, for example, mental states would simply be brain states or brain-body states, making personal persistence ultimately a matter of brain persistence. Such a perspective, however, can easily be charged with “species chauvinism” (Block, 1978, p. 265) because it denies that animals, computers or entities with completely different body chemistry can really have minds such as our own. The attractive alternative to this position, known as functionalism, avoids such shortcomings by defining mental states in purely causal terms. This approach is obviously ontologically neutral about what substance instantiates the mental states, a fact which came to be known as the multiple realizability thesis in the scientific literature (Putnam, 1967).

Coming now back to our main concern, reformulating the psychological continuity view in functionalist terms means that a person will persist over an interval as long as its mental states respect a functional-appropriate-successor relation in that interval (see Davis, 1998). Having reached this point it is natural to ask how a functional-appropriate-successor relation should be understood. Lawrence H. Davis, following a remark by Sydney Shoemaker, considers this to mean that a mental state token is caused by some other mental state tokens through the operation of a mechanism and all mental states count as mental states in that particular system according to their functional definition (Davis, 1998, p. 783). The immediate consequence of this formulation is that a number of intuitive cases in which we would be tempted to say that the person persists, viz. brain transfer, would not actually be person preserving if the “naked” brain does not have mental states (be they conscious or unconscious). The second consequences is that the previous position raises a number of interesting questions regarding the personhood of composite systems in which the mental states are instantiated on different architectures, but nonetheless form a causal network. These considerations will be the subject of the next section.

3. The Extended mind and the extended self

The extended mind thesis formulated in the late 90’s by Andy Clark and David Chalmers (Clark & Chalmers, 1998) has generated intense debates in both philosophy and cognitive science. Without having here the necessary space to offer an elaborate exposition, we will refer only to the main principles that have importance for the present inquiry. At the core of the extended mind thesis, sometimes referred to as active externalism (Clark & Chalmers, 1998, p.8), is the so-called parity principle. This principle states that “if, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is...part of the cognitive process” (Clark & Chalmers, 1998, p.8; see also Clark, 2008, p.222; italics in the original). To support this remark, Clark & Chalmers ask us to imagine a case in which two people have the desire to visit a museum, but one of them suffers from a debilitating case of dementia (Alzheimer’s disease in the original Gedankenexperiment). While the first person has no problem in retrieving from memory the address of the museum, the second one has to rely on a notebook on which the address is written to get to the destination. Thus, as Clark concludes “his notebook plays the role usually played by a
“biological memory” (Clark, 2008, p. 227), providing that the man-artefact coupling relation satisfies the following four criteria: (i) the agent has constant access to the resource; (ii) the information stored is easily and directly available; (iii) the agent automatically endorses the information retrieved from the resource; and (iv) the contents stored in the resource have been endorsed by the agent in the past and are there for this reason (Clark & Chalmers, 1998, p. 17).

To get from here to the “broader view, and see agents themselves as spread into the world” (Clark & Chalmers, 1998, p. 18) might seem a step too far. We will not pursue here the line of argument advanced by Adams and Aizawa that such provisions rest on a fallacy of not distinguishing between the causes and the constituents of a cognitive process (Adams & Aizawa, 2008, p. 83) because the extended mind thesis is not necessarily committed to the idea that external tools are themselves subjects of cognitive processes. A better argument would be to say that an external device could not actually store a person’s beliefs because the stored statements do not have intrinsic intentionality (Adams & Aizawa, 2001, p. 48). Indeed, whether one adopts representationalism, viz. the thesis that cognitive processes involve representations having original content, or the stronger thesis known as the language of thought hypothesis (Fodor, 1975), it appears that the propositional content of a person’s belief stored on an artefact derives its representational character from conventions and not through causal interactions with the world. There are, of course, other arguments against Clark and Chalmers’ original view, but we will come back to them at the end of this paper.

An interesting example of why the extended mind thesis leads to strange and problematic consequences if taken to be true was given by Brie Gertler (2007) in an attempt to “narrow” the concept of the mind. Gertler starts from the simple idea that beliefs and desires generate volitions, which in turn, ceteris paribus, lead to actions. Granting that Clark & Chalmers premises are all correct and the dispositional or non-occurrent beliefs of a pathologically memory impaired person are simply stored on an external device as opposed to his biological memory, leads to the conclusion that his actions are causally determined by an external artefact. The problem appears, says Gertler, when not only dispositional beliefs, but also desires are stored on an external device and that device is not a simple artefact, but an agent-like robot (Gertler, 2007, pp. 196-197). Now, if instead of a single robot, the person stores his standing beliefs and desires on the memory chips of an army of robots, this would amount to a classic case of “branching” or fission that violates synchronic qualitative identity conditions (Gertler, 2007, p. 197). Thinking that the consequences are absurd in such a case, Gertler then moves to advance her own thesis according to which the mind of a person should be restricted only to the phenomenally conscious tip of the iceberg. Like Lawrence Shapiro, we consider this response to be “trading one absurdity for another” (Shapiro, 2011, p. 198), mainly because it severely restricts the psychological continuity criteria for persons in a manner that is simply uncalled for. Of course, Gertler does offer a plausible solution to this situation, saying that even in this case, all mental states of a person are casually linked to dispositional, abyssal structures of the psyche. However, this is not the only possible refutation that can be given to the extended mind thesis.

From a more pragmatic and empirical point of view, we consider that not even biological memory respects the four criteria advanced by Clark and Chalmers (1998, p.17). As Kourken Michaelian observed, biological memory is deeply unreliable, constructive, selective and contextual (Michaelian, 2012). Ergo, external memory is not really memory in the natural sense. Having reached this point it might be asked, however: how about neuroprosthetics – do they enhance in any way the boundaries of the mind and the self? Here, we believe, the answer will be ambiguous at the moment. Devices such as retinal projectors simply generate images on the retina but do not actually process them like the visual human module does. As such, they provide extra input, but not extra processing, making projectors causes rather than constituents of perception. Intracortical visual prosthetics, on the other hand, veer closer to what we might call constituents of cognitive mental states, but even if such devices will finally be made, they do not replace cortical processing in any way, nor are they integrated directly to other cortical areas relevant for conscious visual perception. Integrating artificial tools with the working brain, we believe, would ultimately amount to breaking the so-called “neural code”.

4. Conclusions

I’ve started this paper emphasising the fact that the neo-Lockean psychological continuity thesis still remains today the most popular and plausible conception over what counts for the persistence of persons. From a
functionalist point of view, this amounts to saying that the mental states of P’ at t’ are causally linked to those of P at t through an underlying appropriate mechanism in order for P and P’ to be the same person. The multiple instantiations perspective, however, is silent on the ontological substrate of mental states, making it possible for persons to survive inter-platform transitions as long as these transitions are functionally reliable. Adopting a computational theory of the mind, active externalists took this functionalist view further, maintaining that some parts of the human mind already extend outside of the physical body, forming coupled cognitive systems with tools or other artefacts. Following Adams & Aizawa (2001) we consider “contingent transcranialism” to be perfectly possible but empirically inexistent at present time as we have already argued. Still, as Ned Block (1978) and many other philosophers have remarked over the time, functionalism really does leave outside the famous “what is like to be” (Nagel, 1974) character of phenomenal consciousness. Perhaps the multiple instantiations thesis is not as strong as it was once thought 40 years ago and our sense of the self is inextricably linked to some intrinsic qualitative character of mental states that cannot appear in inorganic substances. If this consideration is taken to be true, in pure Parfitian fashion (Parfit, 1984), you would survive a transition, but it will not be “you”.

References