

# Neurotechnology, Invasiveness and the Extended Mind

Tom Buller

Received: 29 November 2010 / Accepted: 5 August 2011 / Published online: 18 August 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** According to a standard view, the physical boundary of the person—the skin-and-skull boundary—matters morally because this boundary delineates between where the person begins and the world ends. On the basis of this view we make a distinction between invasive interventions that penetrate this boundary and non-invasive interventions that do not. The development of neuroprosthetics, however, raises questions about the significance of this boundary and the relationship between person and body. In particular it has been argued by appeal to the Extended Mind thesis that mind and person can extend beyond the body, and hence the skin-and-skull boundary is of questionable significance. In this paper I argue that the Extended Mind thesis is consistent with the ethical relevance of the skin-and-skull barrier. Although it can be argued that cognitive processes and aspect of mind can extend beyond the skin-and-skull boundary as EM claims, it does not follow that the person is also extended beyond this boundary. The moral sense of person is closely related to the notion of person as a subject of experiences and this, in turn, is related to the sensory and somatosensory aspects of the body. The development of neuroprosthetics provides us with reason to see that persons can be variously embodied, but this is

consistent with the functional and ethical significance of the skin-and-skull boundary.

**Keywords** Neurotechnology · “Extended Mind” · Neuroethics · Invasiveness

## Introduction

If a child were to ask the question, “Where do I begin?” a likely answer from a suitably proud parent would be to say, “At the skin” or to tap the child on the arm. This sort of response reflects a commonly held view that the limit of the person, the line where the person ends and the world begins, occurs at the physical boundary of the body. In this paper I wish to examine the stability and relevance of this boundary in setting these limits. One reason to think that this boundary may not be as stable as first appears follows from consideration of brain-machine interfaces (BMI’s). For BMI’s provide examples of how persons can be variously embodied, as being composed of both biological and artificial parts. It is customary, perhaps, to think of BMI’s as devices, that is to say as “mere” tools that are fundamentally independent of the body. Accordingly, we think of a cochlear implant in broad terms as we do a hearing aid. But as these “tools” become more integrated into the phenomenal and functional world of the person, the more that we might be led to think of these devices as part of the body, and hence as part of the person.

---

T. Buller (✉)  
Philosophy Department, University of Alaska Anchorage,  
3211 Providence Drive,  
Anchorage, AK 99508, USA  
e-mail: aftgb@uaa.alaska.edu

But this leads to a further and more challenging question, namely whether person and body must always coincide. Traditionally we think of the skin-and-skull boundary as defining the boundary of the person, but BMI's suggest that it is possible for the person to extend beyond the body, for external elements to be included into the physical and functional world of the person. Support for this position comes from advocates of the "Extended Mind" (EM) thesis. According to this thesis, the mind can be extended into the world and hence we should be prepared to grant that persons can be extended, as complex, coupled systems unlimited by the skin-and-skull boundary.

In this paper I wish to defend the significance of the skin-and-skull boundary. The argument that I present is that this boundary can itself be understood in functional terms and variously realized, and hence it is consistent with the development of BMI's and the Extended Mind (EM) thesis. As a consequence of this, I believe that there are good reasons to continue to maintain that person and body coincide, and that invasions of the body are *prima facie* metaphysically and ethically significant. In the first section of the paper I present a brief discussion of BMI's focusing on those devices that are specifically designed to replace lost motor function, as well as considering the notion of invasiveness. In the second section I turn to an examination of an argument presented by Joel Anderson [1]. Anderson argues that the skin-and-skull boundary lacks metaphysical and ethical significance and hence the "Invasiveness Criterion"—the criterion that physical invasions matter morally—should be rejected. According to Anderson the EM thesis supports such a conclusion; however I hope in the third section of the paper to show that this is not the case. In the fourth section I discuss a second aspect of Anderson's argument that pertains to the lack of parallel between the notions of inside the body and "inside" the person. I attempt to show that this difference is consistent with our common perspective of the coincidence of person and body, and that the notion of "inside" the person being employed is not supported by the EM thesis. In regard to this latter thesis I argue that the sensory and somatosensory aspects of the body are core elements in our notion of personhood, and thus there are strong reasons to claim that person and body must coincide even if aspects of the mind can be off-loaded into the world.

## Devices and Invasions

In considering the merit of one type of medical intervention over another it is common to regard the invasiveness of the intervention as morally relevant. For example, it would seem clear that in a choice between two types of surgical intervention to repair damaged knee ligaments we should choose the less invasive, other things being equal. The same consideration should lead us to recommend the medication for asthma that has fewer adverse side-effects than the one that has more, and to support the policy that we should use bed-restraints on a patient only in exceptional circumstances. In general terms one can say that an intervention is invasive if it causes the person pain and suffering, or renders her more vulnerable to further injury, or restricts her freedom or autonomy. In describing an action or intervention as invasive it is not necessary that the intervention be physical, as the example of an invasion of privacy reveals; however, perhaps the most common understanding of invasiveness is in physical terms, as pertaining to interventions *into* or involving the body. According to this understanding, a necessary condition of an intervention being invasive is that the intervention involves penetration of the skin-and-skull boundary. This penetration could involve the literal penetration of this boundary, for example, a needle-stick, or an intervention like a gamma-knife that penetrates the body without being in direct contact with it.

It seems plausible to claim that what underlies the ethical relevance that we place on invasions of the body is that we assume a person to be coincidental or coextensive with his or her body, that is to say, we think the skin-and-skull boundary to be significant because it demarcates the line between where the person begins and the world ends. To put the matter in simple and graphic terms: if you stab my arm then you are stabbing *me*; or to put the matter somewhat more precisely: invasions of the body are generally thought to be sufficient to be invasions of the person, and invasions of the person are generally thought to be necessary to be invasions of the body. These claims hold only generally because there can be physical invasions that are not personal invasions, for example, in cases where there has been a loss of personhood; and there can be invasions of the person that are not physical invasions, for example when a person might

feel “invaded” by receiving insults or threats. Furthermore, one should also make a distinction between “macro-level” invasions like pushes or shoves and the “micro-level” invasions of viruses and bacteria, for we would regard micro-level interventions to be invasive only if they were to lead to impairments in the person’s health. Additionally, there are invasions that are physical but their invasiveness can be understood in psychological terms. For example, one might object to the use of SSRI’s on the grounds that they pose a threat to a person’s “authentic” self [2, 3]. Although psychopharmacological interventions effect change through their neurophysiological properties, we think of them as psychological rather than physical invasions.

Advances in neuroscience raise questions about the coincidence of person and body, and, in particular, about the significance of the skin-and-skull boundary. First, the neuroimaging of patients in minimally conscious states provides evidence that some of these patients may, in fact, have higher cognitive capacities that was previously thought. For example, in a recent study involving 54 minimally conscious patients, Monti et al. found that five were able to willfully modulate brain activity. Similar findings were reported by Schiff et al. [4, 5]. These types of findings raise, perhaps, the rather disturbing possibility that not only might our prior diagnosis be mistaken but also, more importantly, that a person may be “in there,” trapped inside the body unable to move like Jean-Dominique in *The Diving Bell and the Butterfly* [6]. In the case of a “locked-in” individual like Bauby we would seem to have a case of an individual who almost entirely lacks the ability to move but whose cognitive and other neurological functions are mostly intact [7]. Since Bauby retains this high degree of cognitive function and the capacities for sentience and self-consciousness, we might conclude that there is good reason to claim that he is still a person. Relatedly, we might view this case as a case of extreme physical disability, and point to a common intuition that reduction in physical ability does not warrant any reduction in the ascription of personhood. This suggests that the physical boundary of the body has questionable moral and metaphysical significance, for it would seem clear that in this case there could be physical invasions that are not invasions of the person.

Moreover, the very notion of a person being trapped “inside” the body implies that the body is not definitive of the person.

A different conclusion can be drawn if our emphasis is on sentience and the capacity for action. For the sake of argument, let us imagine the case of an individual whose physical limitations exhaustively circumscribe the ability to act. If we regard intentional *action* as a necessary condition of personhood, rather than the capacity to have the intention to act, then we might conclude that such an individual fails to meet the conditions of personhood. A similar conclusion could be drawn if one imagines a case in which the individual not only lacks motor function but there is also a severe deficit in somatosensory function. Here we might conclude that since the person is only minimally sentient the conditions of person are not met. Or one might draw the same conclusion on the grounds that the individual is not “embodied” in a robust functional sense since he or she lacks any sensory and somatosensory relationship to the world. In other words, persons cannot be simply “thinking things.”

Brain-machine interfaces (BMI’s) raise a different set of questions about the relationship between person and body [8, 9]. The term “brain-machine interface” (BMI) tends to conjure up notions of cyborgs, androids and other types of “bionic” entity. We can certainly imagine novel devices that enhance our visual or auditory acuity, or our cognitive or physical abilities, that are unlikely to be developed anytime soon; however, various types of BMI have been in use for over forty years and are currently used to treat a variety of conditions. These include the use of cochlear implants to restore lost hearing, spinal cord stimulation for the treatment of pain and implantable devices for the treatment of Parkinson’s disease [10, p. 138]. In addition, there has been progress in the development of BMI’s that are designed to replace lost motor function to individuals who have suffered paralysis as a result of severe spinal cord injury [11]. As these various technologies continue to develop one can expect that so also will the range and level of these interactions, and future neuro-technologies may provide a level of interaction that compensates more fully for the loss of motor function. Furthermore, research has also been conducted on devices that can substitute one sensor modality to another, for example, how a person might learn to “see” through tactile information [12].

As Donoghue states [13] there is, as yet, no standard vocabulary regarding BMI's and different terms have been used including "neural interface systems," "brain computer interfaces" and "neuro-prosthetics." However the term "neuroprosthetic" is often taken to refer to a subset of BMI's, namely those that are designed to replace lost motor, as opposed to sensory or cognitive function. In this paper I will use the term in this more limited sense since it suggests how a BMI can serve as a functional interface between brain and world, and restore or enhance connections, or even create new connections, between the person and the environment. However, this usage is not meant to exclude devices that are designed to restore (or enhance) sensory function or cognitive function, as opposed to motor function.

Neuroprosthetics designed to replace lost motor function detect the neural signals that correspond to the individual's goals or intentions, and then transform the signals into commands by communicating this information to an external, prosthetic device. Neuroprosthetic BMI's can be differentiated in a number of ways, however a familiar distinction is between invasive devices that are implanted either subdurally or intracortically, and non-invasive devices that are placed on the surface of the head. Invasive devices detect either action potentials or field potentials, and can record neural signals of high quality from single or multiple recording sites. Non-invasive devices detect potentials from large neuronal populations and record signals of lower quality, although this is sufficient to detect brain activity that correlates with voluntary intention, gaze angle, cognitive state and visual stimuli [13 p. 537]. Since invasive devices are implanted they are more suitable for long-term use than non-invasive devices that are attached to the scalp.

A further advantage of invasive devices is that they can "map neural activity related to the intended motor feature directly to the desired action" [13, p. 513]. Accordingly, such devices are direct systems and are distinguished from invasive, indirect devices in which the desired action is mapped to surrogate neural activity, for example, suppressed cortical rhythms [13, p. 512]. The advantage relates to the extent to which the neuroprosthetic device is "transparent." A device is transparent if the intended action, for example, the movement of a prosthetic limb, follows directly from the relevant mental state, namely the intention to

move the limb, rather than from brain activity that has been as associated with the action through learning. For the most part our brains and bodies are direct systems because it is generally the case that the voluntary movement of the body follows directly from the relevant intention to act. (If it were the case that in order to move my arm I would have to learn, for instance, to suppress the cortical rhythms in my brain, or even to consciously think about moving my arm, the action would take a lot more work, at least initially). It is plausible to contend that the more transparent that a device is, that is to say, the more that the relevant action follows directly and economically from the intention to perform that action, the greater the chance of neuroprosthetic success.

A related way of thinking about the transparency of a device is in terms of the degree to which the person using the device becomes unaware of its intermediary role. For example, a person who has recently lost her sight might learn to navigate the world through the aid of a cane. In learning to use the cane one can imagine that initially the feeling of the cane dominates, and the prevailing sensation is that of feeling the world through the cane. But as the person becomes more familiar with its use the more the cane becomes part of the person's own perceptual system. This is not to suggest that the person becomes unaware of the tactile sensation but that these integrated sensations are what it means to perceive the world. In this regard, we can regard the cane as a form of sensory substitution in which touch and sound are substituted for vision (or more radically, perhaps, one that creates a different form of vision). Thus it would equally inappropriate to say that the cane is a device through which the sight-challenged person perceives the world as to say that the eyes are a device through which a sighted person sees the world. For properly understood, these are not devices but parts of the integrated coupled system.

This distinction between a "mere" device and an integrated system relates importantly to the notions of person and embodiment, and the relevance of the skin-and-skull boundary. For example, consider the case of an individual with substantial motor limitations who is able to control a robotic arm with the aid of a neural implant [14, 15]. In such a case we can think of the "boundary" of the person in a number of different ways. First, one might hold what might be termed a "naturalistic" view according to

which a person is thought to coincide with the original body and, therefore, the biological skin-and-skull boundary. Accordingly, since the robotic arm is artificial—a prosthetic—it is beyond this boundary and hence not part of the person. Consequently it should be regarded as a “mere” device. Furthermore, if one holds the view that an intervention is invasive only if it invades the person, then an intervention into or of the device would not, strictly speaking, be invasive.

Second, one can think of the boundary of a person and the body in functional rather than in physical or biological terms. In this case something legitimately counts as part of the body if it plays the appropriate functional role, rather than whether it is biological or artificial, natural or novel. The robotic arm should, therefore, be regarded as being part of the body to the extent that it plays the same or similar functions as the original, biological arm. Since we are still early on in the development of such neuroprosthetic devices it is true that these devices are limited in their functional role; however, in the future case of a robotic arm that is functionally and phenomenologically identical to the original biological arm, it would seem difficult to resist the conclusion that the neuroprosthetic should be regarded as being part of the body and person. In a parallel fashion one can argue that what matters in determining whether a heart valve or replacement knee is part of the body is its functional role, not the type of material from which it is made or its origin. This functionalist view thus supports the notion that the body can be variously realized, and that a device should be regarded as part of the body according to the extent to that it is part of the integrated system. Interestingly, as these technologies develop further we may come to rethink the very idea of a “prosthetic” for we generally use this term to refer to something that is novel and artificial. As we become more familiar with the “plastic” nature of the body, the less relevant will become the question of its specific physical constitution.

The point could be raised that the skin-and-skull boundary retains its moral and metaphysical significance even if we adopt this more functionalist view, for this boundary still marks the line between where the person begins and the world ends. In this regard, one might be happy to grant that what matters is functional integration rather than physical nature, but maintain, nevertheless, that part and parcel of this functional integration requires consistency with this

boundary: an “artificial” skin must perform the same function as natural skin. In response one might claim that the skin-and-skull boundary itself can be understood in functional terms and variously realized. For instance, to return to the case of the sight-challenged person discussed above, one might regard the cane as part of the body and person even though the cane provides a way of seeing the world that is functionally and phenomenological novel.

The functionalist approach regards a device as part of the body according to the extent that it plays the appropriate functional role, and if we regard person and body as coincidental then this would give us grounds to regard the device as being part of the person. A third way in which we might view a neuroprosthetic suggests that a device could be part of the person but not part of the body. In considering the robotic limb one might conclude, as the “naturalist” view above does, that the device is not part of the body, and it would seem to be obvious, therefore, that the device could not be part of the person. But is this correct? Intuitively it seems odd to claim that something could be part of the person but not part of the body, for our common intuition leads us to regard the skin-and-skull boundary as marking the boundary of both person and body; in other words, the notion of a person extending beyond the body appears implausible. But perhaps this is just a prejudice. As the Extended Mind thesis proposes, if the argument can be made that part of the world plays the same role as traditionally played by something within the skin-and-skull boundary, then perhaps we should question our assumption that persons and bodies must coincide [16]. And if this is correct, then the ethical and metaphysical relevance of the skin-and-skull boundary would seem to be in doubt.

### “Bad Metaphysics”

Joel Anderson has argued that there are “deep conceptual and ethical problems” with what he terms the “Invasiveness Criterion” [1]. The criterion is described as “the idea that what makes some neuroprosthetic technologies problematic is that they violate a boundary between what is inside the person and what is outside” [1, p. 259]. According to Anderson, the problem with this criterion and the supposed ethical relevance of the skin-and-skull



boundary is that it rests upon what he terms “bad metaphysics.” There are, he believes, “good reasons to question the idea that my *person*, as something not to be transgressed or violated, coincides with my *body*” [1, p. 264]. Anderson presents two main arguments in support of his position. The first of these arguments appeals to the “Extended Mind” thesis and the second to the notion of functional and phenomenological transparency.

The Extended Mind (EM) thesis claims that the mind itself need not be “in the head” but can extend into and incorporate parts of the world [16]. In conjunction with the internal, neural states, these external elements form part of a “coupled system” in which cognition is distributed. In broad terms, EM can be understood as a functionalist thesis, one that maintains that the physical realizers or vehicles of cognition can lie outside the skin-and-skull boundary. More specifically it is the view that the material aspects of a system that allow the system to have content-full mental states—the vehicles of these states—can be “distributed across brain, body, and world” [17, p. 1].

As Andy Clark has stated, the original argument for EM took the form of a series thought-experiments that sought to challenge the traditional internalist view that the mind is confined to the head [17 p. 1]. One of these thought-experiments describes a person, Otto, who suffers from a mild form of Alzheimer’s disease and who carries a notebook in which he writes down information. Otto hears that there is a new exhibition at the Museum of Modern Art (MOMA) and consults his notebook to retrieve the museum’s address—53rd Street. According to EM, Otto has the belief that that the museum is on 53rd Street *prior* to his consulting the notebook, since the notebook plays the same functional role for Otto as biological memory does for a “normal” person. This functional role is defined in terms of four key features: the notebook is a constant in Otto’s life; the information contained therein is reliable; the information is automatically endorsed by Otto; and the information has previously been consciously endorsed. [16, p. 17]. As Clark and Chalmers contend, if we accept the functional similarity between Otto and the standard case then we should be prepared to endorse the following “Parity Principle.”

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 accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process. [16, p. 8].

As Clark says

In other words, for the purposes of identifying the material vehicles of cognitive processes, we should (normatively speaking) ignore the old metabolic boundaries of skin and skull, and attend to the computational and functional organization of the problem-solving whole [17, p. 2].

Anderson’s second argument against the criterion pertains to the notions of experiential transparency and performance. He presents two versions of a case of auditory enhancement: “BCI-Anna” who has a neuroprosthetic device implanted into her brain to enable to hear and to discriminate musical sounds in a superior way, and “PDA-Anna” who has a similar device attached externally, for example, to her belt. Importantly, the devices are identical in terms of experiential transparency and performance, that is to say, if we swapped one device for the other then we can suppose that neither PDA-Anna nor BCI-Anna would notice any difference. Anderson draws the following conclusion from this case.

The point for our present purposes is that once we have the cases in which there are no differences in performance or experiential transparency between BCI-Anna [and] PDA-Anna... then it is unclear what non-question-begging argument there could be for saying how BMI-Anna’s implant is “inside” her person in a way that PDA-Anna’s handheld device is not. But then, once it becomes clear that the distinction between inside and outside the body does not coincide with the distinction between inside and outside the person, it becomes unclear why the boundary of the skin should be seen as ethically significant [1, p. 266].

At first glance it would appear that Anderson is right to claim that there are “good reasons” for thinking that “my person” and “my body” do not coincide. For it would seem to be clear that if the mind extends beyond the skin-and-skull boundary, then so does the person. Furthermore, if phenomenal and functional role is essential for a device being “inside” rather than its physical location, then we

should not think of invasiveness in terms of invasions of the physical boundary. However, despite their initial appeal, I hope to show that Anderson's two arguments are not decisive. Specifically, I argue that: first, EM is consistent with the criterion; second, the conditions of equality and experiential transparency that underlie the case of BCI and PDA-Anna are not supported by EM; and third, this case is consistent with the notion of persons being variously embodied.

As a preliminary step, it is helpful to consider what it might mean to say that "my person" and "my body" do not *coincide*. Arguably, no concept has provoked as much discussion and controversy in bioethics over the last few years as the concept of person. This controversy is generated in considerable part by the plurality of views regarding the necessary and sufficient conditions of personhood, as well as the theoretically decisive role that the concept has played in resolving ethical questions. For many, questions about the morality of abortion, or the moral status of animals or the permanently comatose can be resolved or, at least, greatly benefited, by determining whether or not the conditions of personhood are met.

In our attempts to explicate the concept of person the distinction has been made between the "metaphysical" and "moral" senses of the term. Briefly, when considered in its metaphysical sense, the focus of attention is typically on questions regarding the relationship among person, self, and body, and the necessary and sufficient conditions of personhood at one time and over time. In contrast, when considered in its moral sense discussion tends to focus on those properties that are thought to uniquely distinguish persons as the objects of the highest moral concern and status. As we are familiar, putative candidates for these properties include sentience, self-consciousness, level of organization, or and the potential for the possession of these capacities. In some cases these two senses are related, for instance, in discussions about praise and blame. For it is generally thought that it is appropriate to praise or blame a person for an action only if he or she is the person who performed the action. In other cases, however, the two senses are unrelated. For example, in attempting to determine whether an individual or group or animal should be appropriately judged to be a person we are generally unconcerned with the question of whether "person" is a metaphysically robust notion. By this I mean that we are prepared to consider the question whether, for

instance, a permanently comatose individual is a person without having to commit to the view that "person" is a natural kind or even a coherent type of artifact kind. This separation of the moral and metaphysical senses can also be seen in another context, namely the debate about the implications that advances in neuroscience will have on the law. The argument has been made that advances in neuroscience present no challenge to the law and determinations of responsibility because the law is not interested in free will in any "metaphysical" sense. Accordingly, one might say that the law is merely interested in the moral sense of free will.

According to his definition, Anderson defines "my person" in its moral sense as "something not to be transgressed or violated" [1, p. 264]. The central claim against the Invasiveness Criterion is that the criterion rests upon "bad metaphysics," and this suggests that the criterion would be in better stead if its metaphysical grounding were more secure. But as the above suggests, one might argue that the concept of person and terms such as "my person" need not be metaphysically robust. When one says that a thought is "inside" the head we think it appropriate to use quotation marks because we do not normally mean to be understood as making a claim about the specific location of the thought. In a parallel fashion we could say that in claiming that an experience occurs "inside" the person we are making no commitment as to the actual location of the experience. If this is correct, then the conclusion that "my person" and "my body" do not coincide might not be too telling.

Underlying the claim the claim that "my person" and "my body" do not coincide rests an understanding of "my person" in first-personal, functional and phenomenological terms as the subject and object of self-interested concern. A person is concerned for the welfare of his or her person, and this is because "my person" coincides with *me*—a violation or transgression of my person is necessarily a violation or transgression of me. But it seems clear that a person need not be concerned for the welfare of his or her body, as Anderson's example of his indifference to the welfare of the fingernails he is about to cut off suggests. Rather, we are concerned for the welfare of the body only because invasions of the body are invasions of the person: by invading my body you are invading me. "My person" is invaded if you cause me pain and penetration of the skin-and-skull boundary is usually

sufficient for this to occur; conversely, if this boundary is penetrated but there is no adverse functional or phenomenological effect, then we would have reason to doubt that the person has been invaded. The case of Otto suggests that there can be functional invasions of the person that are not physical invasions of the body. In this case one could argue that Otto-and-the-notebook—Otto *qua* person—has been invaded even though the skin-and-skull boundary, and hence his body, have not. This provides reason to suggest, therefore, that an invasion of the body is neither strictly necessary nor sufficient to be an invasion of the person. Hence, “my person” and “my body” do not coincide.

### Extended Minds and Persons

We can now turn to the first of Anderson’s main arguments and his appeal to the Extended Mind (EM) thesis. For the sake of argument, let us accept EM and grant that the mind itself can be off-loaded into the world. We can also grant that Otto has the dispositional belief that MOMA is on 53rd Street and that Otto-and-the-notebook is a coupled system. Thus we have accepted that the vehicles of cognition can be extended beyond the skin-and-skull boundary and incorporate elements in the world. In Otto’s case this means that we are prepared to grant that Otto’s mind is extended and that the information contained in the notebook is part of his mind.

The crucial question to ask is whether by granting these claims we have thereby granted that there are good reasons to think that person and body do not coincide. At first glance it would seem obvious that by extending Otto’s mind we have thereby extended Otto *qua* person. For if we are prepared to think of Otto as a coupled, cognitive system and to hold that the mind can be extended into the world, then it would seem to be a short step to regard this extended, coupled system as coincidental with the person. Presumably, the argument for such a conclusion is something along the line of the *Parity Thesis*: if we are prepared to grant that Otto is a person prior to the onset of Alzheimer’s, then we must conclude that Otto-and-the-notebook is also a person, for there is no relevant functional difference between the two of them. If this is correct, then we would seem to have good reason to reject the coincidence of person and

body, since here we have a case where the person extends beyond the skin-and-skull boundary.

Nevertheless, one can respond to this line of argument in the following way. Although adoption of EM provides us with reason to conclude that the mind can extend beyond the skin-and-skull boundary, it warrants the further conclusion that the person can extend beyond the body only if one accepts the further claim that person and mind coincide. To put the point slightly differently: acceptance of EM would be sufficient to show that “my person” and “my body” do not coincide if one also accepted the additional claim that “my person” and “my mind” always do coincide. For in this case we would have accepted that the mind can extend beyond the body and that the mind and person coincide. But do we necessarily extend the person by extending the mind? There are a number of considerations to suggest that this might not be the case.

First, the *Parity Thesis* itself might suggest a reason to think that mind and person can “come apart.” In abstract terms one can think of the *Parity Thesis* as a thesis about the location of cognitive processes. The thesis suggests that if two sets of processes play the same functional role, then it should not matter whether the process is wholly or partially internal. One can put the point another way by saying that there is no requirement that cognitive processes are *necessarily* internal—they can be located either inside or outside the body, or both. A parallel application of the *Thesis* could be used to show that cognitive processes need not be necessarily internal to the person, or to put the matter differently, a process is internal to the person simply in terms of its functional role and organization. Thus if we return to Otto one might contend that this case shows that the mind can be extended beyond-the-skin and skull boundary, as Anderson claims; but one might equally contend that the case shows that the mind can extend beyond the person and that persons can be embodied individuals with extended minds. In other words, Otto *qua* person coincides with the skin-and-skull boundary even though his mind is extended beyond this boundary.

It might be objected that it is very odd, indeed, to think of the mind as being “outside” the person, not least because there is substantial support for the view that person and mind are intimately and essentially related, and that loss or substantial deterioration of the mind implies the destruction of the person. But



perhaps there is no greater metaphysical need that *all* aspects of mind occur in the person as that they *all* occur in the head. Accordingly, we can describe Otto as a person who has off-loaded some elements of his mind, namely some of his beliefs, and stored them in a device outside of his body. These off-loaded elements can be considered to be outside of his person on the grounds that they are minimally integrated or minimally part of the coupled system. The point being made here is not that persons can be disembodied but rather that there could be embodied aspects of a person's mind that are not part of the person. Thus Otto-and-his-notebook should be regarded as a coupled system in which the mind extends beyond the skin-and-skull boundary, but Otto *qua* person, as "something not to be transgressed or violated," however, is still demarcated by this boundary.

As mentioned above, it is certainly the case that we routinely think of mind and person as coinciding and that there is considerable support for the view that the identity of the person is a matter of psychological continuity. Deterioration or substantial destruction of the cognitive processes is thus thought to entail the loss of personhood and the elements that are fundamental to constituting the identity of the person. This understanding of the relation between mind and person is not at odds with the notion that aspects of mind can exist beyond the person, however. In Otto's case we can describe the case as one in which the parts of his brain that realize the psychological elements that are constitutive of Otto are intact, but a number of processes that are minimally integrated into the overall cognitive system have been off-loaded. We can, therefore, describe this case in at least two ways: first, as one in which certain cognitive processes have been off-loaded beyond the skin-and-skull boundary but remain part of the person; or second, as one in which these processes are outside this boundary and the person.

Further support for the notion that "my person" and "my mind" may not coincide may be gained by considering the aspects of mind that can be "off-loaded" as proposed by EM. One of the core elements of our standard notion of person is that persons are the subjects of self-interested concern. Part and parcel of this element is its phenomenal aspect—a person has experiences—and, generally speaking, a person recognizes that these experiences are *her* experiences

and that these experiences are happening to *her*. (This is not to claim that only persons have experiences). EM claims that certain aspects of the mind can be off-loaded into the world, for example, beliefs; EM does not claim, however, that every type of process can be off-loaded.

In granting that the case of Otto is a *bona fide* example of extended mind, one has accepted that certain dispositional, informational states can be determined by external, as opposed to internal, factors. As Clark and Chalmers state the matter, other types of mental state such as experiences, may be determined only internally rather than (partially) externally, and hence resist off-loading into the world [16, p. 12]. One reason for this resistance may be the difficulty in accounting for the phenomenal aspect of experience in informational and dispositional terms. According to Clark and Chalmers, the features of the notebook that play a "crucial role" in satisfying the criterion that some of Otto's beliefs can be off-loaded are that the notebook is portable and reliable: Otto rarely takes action without it, the information is automatically endorsed upon retrieval, and the notebook is a constant in his life. Furthermore, the information contained therein has been consciously endorsed [16, p. 17]. To some degree one might think these criteria would equally apply to experiences. One might contend, for instance, that Otto could store the memory of an experience, or perhaps the details—the informational content—of the experience in the notebook, and he would trust and automatically endorse this information. Nevertheless, experiences seem to be unlike beliefs in that the former are not readily thought of as dispositional states: it seems to be a feature of an experience that it be occurrent, that the experience is going on right now and that the person is aware of this experience. In this regard, "having an experience" is unlike "having a belief" since only the former needs to be conscious. (It could be objected that a person can be aware of a state of affairs without having to be conscious of that state of affairs, for example a person who is on "automatic pilot" while driving. Unconscious states, however, are not dispositional states in the sense that the unconscious states are still occurrent).

In arguing that conscious experiences cannot be extended it is important to distinguish between the creation of new coupled systems and the off-loading of experiences into the world. In the case of the

sight-challenged person discussed earlier we may well wish to draw the conclusion that the person-and-cane forms a new coupled system in which the cane becomes part of the integrated physical and phenomenal whole; in Otto's case, however, there has been an off-loading only of a certain type of dispositional state, namely, belief. If it is plausible to claim that experiences, unlike beliefs, must be occurrent and conscious, then it is not sufficient for experiences to be reliable and trustworthy etc. in order for these states to be off-loaded. Accordingly, we should see Otto-and-the-notebook as a coupled system in which a subset of cognitive processes has been off-loaded that does not include experiences. The challenging question that this raises is whether such a system should be regarded as a person, and the position that I am defending here is that there are at least two possible descriptions of Otto: first, we could describe the case as one in which both mind and person are extended; or second, we could describe it as case of extended mind alone. The second description claims that although Otto-and-the-notebook is a coupled system and an example of extended mind, Otto *qua* person is still defined by the skin-and-skull boundary. On this description, therefore, acceptance of EM does not entail that the person is also extended.

A further consideration in support of this conclusion relates to the moral sense of person and Anderson's definition of the term as "something not to be transgressed or violated." At first glance this definition would appear to be too broad for it includes a number of candidates that we would normally exclude, for example works of art or the environment. To render the definition less liberal I think that it is helpful to understand the term "my person" as being closer to the notion of a subject of self-interested concern. One of the ways in which one might attempt to determine whether a device should be regarded as part of the person is in terms of the level of its integration into the overall phenomenal and functional organization. Moreover, it would seem that the more integrated the device, the more the person will regard invasion or damage to the device as harm to the person herself. Conversely, if the person regards the device instrumentally, as a tool or device that serves to perform a specific function, the less the

person will regard damage to the device as damage to herself. (This is not to deny that harm to "instrumental" devices cannot be catastrophic to the person). In Otto's case it is certainly true that the loss of the notebook may be as catastrophic as the further deterioration of the diseased neurons in his head, but the simple degree of harm may not be sufficient for Otto to regard the notebook as part of his person. How Otto might regard the notebook is, of course, not definitive in determining whether or not there are good grounds for thinking that Otto-and-the-notebook is a person, and Otto might regard the loss of the notebook as loss or damage as harm to himself. But this type of self-interested concern is related to functional and phenomenal integration, and hence may be relevant to determining the boundaries of the person.

Although one might agree that the notebook should be considered to be part of a coupled system, the notebook lacks the sensory or somatosensory elements that are often important in determining the boundaries of and as being essential to the person. It is true that Otto might regard the loss of the notebook as literally the loss of part of his mind but from this it does not follow that he would regard its loss as damage to *him*. Otto might regard the notebook as an essential but "mere" device to store information: as being a device that is part of the integrated, coupled system but lacking the sensory elements that are often regarded as important in determining the person. In other cases, however, one might draw the conclusion that damage or harm to the prosthetic amounts to harm to the person, as in the case of the individual with a robotic limb or the sight-challenged person with the cane. But here it seems that there is an extension of somatosensory function to the prosthetic limbs, and that there is a higher degree of functional and phenomenal organization.

Otto presents us with a case in which the mind and certain dispositional mental states have been extended beyond the skin-and-skull boundary. In contrast, let's imagine the case of Gretl who differs from Otto in that a greater number of cognitive processes have been extended to her notebook (or whatever other device), including experiences and other conscious mental states. In Gretl's case there would seem to be good reason for thinking that the notebook is just as much part of *her* as any other physical part. And the more the notebook is integrated into a single

functional and phenomenal whole, the more we would be prepared to regard Gretl-and-the-notebook as a single, unified coupled system and subject of experiences. But here again there would appear to be two rival descriptions of the case: on the one hand we could say that this is a robust case of extended mind in which a considerable number of processes have been off-loaded into the world; on the other hand one could claim that this is an example of novel embodiment in which the skin-and-skull boundary has been re-drawn: it is not that the mind has been extended but that Gretl has been re-embodied.

To return to Otto, one way to help with the further deterioration of his brain is to make use of neuro-prostheses and off-load cognitive functions to his notebook or other portable device on a belt around his waist. Furthermore, let us assume that this new system is identical to the old system in terms of performance and phenomenal experience. It seems legitimate to say that we would have thereby moved Otto's brain from inside his skull to around his body, rather than extending his mind into the world. For in this case it seems difficult to point to the external elements that are playing a central role in the cognitive process. Thus the challenge facing the appeal to EM would seem to be this: on the one hand, if only certain dispositional states can be off-loaded then there is reason to think that the person continues to coincide with the subject of experiences as described by the somatosensory boundary of the body; on the other hand, if experiences can be off-loaded and the new device becomes part of the functional and phenomenal whole, then there is reason to regard the new device as part of the body.

The argument has been made that if we adopt EM then we should re-draw our category of person to include societal, environmental and other elements [16–20]. We should, therefore, move away from the notion that persons must be embodied in the traditional sense and be prepared to accept that persons may be complex integrations of individuals and elements in the world. The position that I am defending in this paper is, perhaps regrettably, rather less bold, namely that persons are necessarily embodied even if minds need not be. This view is consistent with the view that persons can be variously embodied and it does not require that the body be a uniform

whole more or less similar to the one traditionally described by the skin-and-skull boundary. Moreover, the view does not require that persons be wholly embodied and that elements of the person cannot extend beyond the body. But this view does maintain that a fundamental way that the body relates to the person is through the body's somatosensory, proprioceptive and sensory functions. These functions enable us to perceive the boundaries of ourselves, to experience ourselves and the world, and to orientate ourselves within the world. It is through these functions that the person is intimately connected with the body and that, in the normal course of events, invasions of my body are invasions of the person.

Accordingly, the sensory aspect of the body can be claimed to be an integral aspect of the person, and thus in the absence of this sensory aspect one might question whether the individual still meets the conditions of personhood. In the present context, the moral sense of person is understood as "something not to be transgressed or violated." In Otto's case one might defend the notion that harm to his notebook constitutes transgression and violation even though the notebook is beyond the skin-and-skull boundary; in other words, this is a case of a personal but not a physical invasion. However, even if we support this conclusion we would be reluctant to draw the more radical notion that person and body do not generally coincide, and the reason for this reluctance is based on the intuition that the sensory aspect of the body is necessary in order for an entity to be "something not to be transgressed or violated." If one removes this sensory element it seems difficult to explain what it is that makes a person an entity that can be transgressed or violated.

As the development of neuroprosthetic devices progresses, we can imagine that future devices will become available that the person is able to control by the relevant intention and which become phenomenally and functionally integrated into the person's perceptual world. Thus we might reach the stage when a robotic limb feels and functions like the original biological one. Once we have reached this level of development it seems difficult to resist the conclusion that the robotic arm should be regarded as part of the body. The skin-and-skull boundary and the other parts of the body continue

to play their prior functional role, and therefore, one could claim that “my person” would continue to coincide with “my body.”

### Experience, Transparency, and Performance

I wish now to briefly consider the second argument that Anderson presents, the case of BCI and PDA-Anna. To recall, BCI-Anna and PDA-Anna both have devices that give them superior abilities to hear musical sounds; they differ, however, in that whereas BCI-Anna’s device is implanted in her head, PDA-Anna’s device is placed externally on her body. The devices are identical in terms of the phenomenal and functional aspects, that is to say, if we were to swap one device for the other neither Anna would be able to tell the difference. According to Anderson, both devices are equally “inside” the person on account of the phenomenal and functional identity; however it is clear that only one of the devices, namely, BCI-Anna’s, is inside the body. Hence a device can be “inside” the person but not inside the body, and this warrants the conclusion that “my person” and “my body” do not coincide.

In considering this argument it seems difficult to resist the conclusion that the devices are equally “inside” the person. For since the devices are functionally and phenomenologically identical it seems difficult to imagine how one could distinguish one from the other. Similarly, if a cochlear implant were to provide a person with a level and type of hearing that was indistinguishable from “normal” hearing and functioned the same way as a “normal” ear, it seems very difficult, indeed, to claim that one device is “inside” the person but the other not. Presumably this is because we understand the notion of “inside” employed here to be simply a matter of phenomenal and functional performance. The important question to ask, however, is whether this obliges us to accept the further conclusion that this case shows that “my person” and “my body” do not coincide, as Anderson claims.

An alternative conclusion to the one that Anderson draws is that this case shows that persons can be variously embodied. As Andy Clark has claimed, one can think of persons as “promiscuously embodied,” that is to say, not limited to any specific type of

realization [20, p. 277]. Furthermore, one need not suppose that the physical realizer be a single physical whole, for there seems to be nothing contradictory in supposing that coupled systems or persons could be composed of a number of integrated but physically distinct parts. Accordingly, one can view this case as one that shows how internal neural processes can be differently realized, rather than showing the lack of coincidence between “my person” and “my body.” In both BCI and PDA-Anna’s cases the brain is a combination of biological and neuroprosthetic elements and the cognitive processes occur in the brain. The cases differ, therefore, only in the way that the processes are realized and not in any deeper metaphysical sense.

The case of Jean-Dominique Bauby shows us that invasions of the body are not necessarily invasions of the person, and the simple physical fact that one has penetrated the skin-and-skull boundary does not mean that the person has been invaded. But this does not compel us to reject the claim that persons are necessarily embodied, and nor does it give us reason to reject the relevance of the skin-and-skull boundary. The case of BCI and PDA-Anna suggests that we should think of the notion of “inside” the person in phenomenal and functional terms, but this is entirely compatible with the idea that these phenomenal and functional terms relate to the fact that the person is embodied.

A final point concerns the differences between the case of PDA and BCI-Anna to Otto. Whereas the former case is described as one in which there is phenomenal and functional identity, in Otto’s case there is no suggestion or requirement that the notebook is phenomenally or functionally transparent. In fact, it is precisely the lack of the requirement that makes the EM thesis such an interesting and challenging one. The appropriate conclusion to draw is that Otto and PDA-Anna are dissimilar and hence the conclusions that we might think are warranted on the basis of EM would not apply in the case of BCI and PDA-Anna.

### Conclusion

If one conceives of persons as being “promiscuously embodied” then one can think of the skin-and-skull boundary itself as being variously realizable. This

boundary need not be thought of in strict physical terms as a single layer or “skin” that wraps the person within but in more functional terms as one of the means by which the person is able to perceive and to move in the world. This understanding of the skin-and-skull boundary provides a way in which to think that “my person” coincides with “my body” and, therefore, a reason why one should retain the “Invasiveness Criterion.” If it is correct to claim that persons are necessarily embodied and that the sensory and somatosensory aspects of the body are core elements of the person, then invasions of the body will generally be invasions of the person.

The development of neuroprosthetics reveals ways in which the body and the person can be variously realized, and as these developments continue we can predict that we will become more familiar with the notion that embodiment can take various forms. In this fashion we are literally and metaphorically reshaping humanity. As this reshaping occurs we are being pressed to rethink what it means to be human and what it means to be a person. The development of neuroprosthetics and the detection of a high level of brain activity in patients previously thought to be “minimally conscious” suggest that the relationship between personhood and embodiment may be yet more complex and controversial than we thought. Nevertheless, if persons are necessarily but variously embodied, and the essential elements of embodiment are the sensory and somatosensory aspects of which the skin-and-skull boundary is a part, then invasions of this boundary, however it is realized, will continue to be ethically relevant.

## References

1. Anderson, J. 2008. Neuro-prosthetics, the extended mind and respect for persons disability. In *The contingent nature of life*, ed. M. Düwell et al., 259–272. Dordrecht: Springer.
2. Parens, E. 2005. Authenticity and ambivalence. *The Hastings Center Report* 35(3): 34–41.
3. DeGrazia, D. 2000. Prozac, enhancement, and self-creation. *The Hastings Center Report* 30(2): 34–40.
4. Monti, M.M., et al. 2010. Willful modulation of brain activity in disorders of consciousness. *NEJM* 362: 579–589.
5. Schiff, N.D., et al. 2005. FMRI reveals large-scale network in minimally conscious patients. *Neurology* 64: 514–523.
6. Bauby, J.-D. 1997. *The diving bell and the butterfly*. New York: Vintage International.
7. Fenton, A., and S. Alpert. 2008. Extending our view on using BCI's for Locked-in-syndrome. *Neuroethics* 1(2): 119–132.
8. Clausen, J. 2008. Moving minds: Ethical aspects of motor neural prostheses. *Biotechnology Journal* 3(12): 1493–1501.
9. Clausen, J. 2009. Man, machine and in between. *Nature* 457: 180–1081.
10. Patil, P.G., and A. Dennis. 2008. The development of brain-machine interface neuroprosthetic devices. *Neurotherapeutics* 5: 137–146.
11. Lebedev, Mikhail A., and Miguel A.L. Nicolelis. 2006. Brain-machine interfaces: Past, present and future. *Trends in Neurosciences* 29: 536–546.
12. Bach-y-Rita, P., and S.W. Kercel. 2003. Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences* 7(12): 541–546.
13. Donoghue, John P. 2008. Bridging the brain to the world: A perspective on neural interface systems. *Neuron* 60: 511–521.
14. Hochberg, Leigh R., et al. 2006. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* 442: 164–171.
15. Anderson, R.A., et al. 2004. Cognitive neural prosthetics. *Trends in Cognitive Sciences* 8: 486–493.
16. Clark, Andy, and David Chalmers. 1998. The Extended Mind. *Analysis* 58: 7–19.
17. Clark, Andy. 2005. Intrinsic content, active memory and the extended mind. *Analysis* 65: 1–11.
18. Nelson, J.L. 2003. *Hippocrates maze: Ethical explorations of the medical labyrinth*. Maryland: Rowan and Littlefield.
19. Levy, Neil. 2007. Rethinking neuroethics in the light of the extended mind thesis. *The American Journal of Bioethics* 7: 3–11.
20. Clark, Andy. 2007. Re-inventing ourselves: The plasticity of embodiment, sensing, and mind. *The Journal of Medicine and Philosophy* 32: 263–282.