



Journal of Cognition and Neuroethics

ISSN: 2166-5087

January, 2016. Volume 3, Issue 4.

Journal of Cognition and Neuroethics

Managing Editor

Jami L. Anderson

Production Editor

Zea Miller

Publication Details

Volume 3, Issue 4 was digitally published in January of 2016 from Flint, Michigan, under ISSN 2166-5087.

© 2016 Center for Cognition and Neuroethics

The *Journal of Cognition and Neuroethics* is produced by the Center for Cognition and Neuroethics. For more on CCN or this journal, please visit cognethic.org.

Center for Cognition and Neuroethics
University of Michigan-Flint
Philosophy Department
544 French Hall
303 East Kearsley Street
Flint, MI 48502-1950

Table of Contents

1	Hemispherectomies and Independently Conscious Brain Regions James Blackmon	1–26
2	The Complexity of Suicide: Review of Recent Neuroscientific Evidence Erica Ching	27–40
3	What is Enhancement? Catherine Gee	41–54
4	Does the Human Right to Health Include a Right to Biomedical Enhancement? Martin Gunderson	55–75
5	Resource Egalitarianism, Prosthetics, and Enhancements Rhonda Martens	77–96
6	Reversibility and Deep Brain Stimulation Jennifer Mundale	97–111
7	A Subjectivist Solution to the Problem of Harm in Genetic Enhancement Sruthi Rothenfluch	113–130

Journal of Cognition and Neuroethics

Hemispherectomies and Independently Conscious Brain Regions

James Blackmon

San Francisco State University

Biography

Much of Blackmon's work focuses on cognition, computation, and consciousness (as he suspects they might have something to do with each other) and on how they could be realized by physical things. He earned his doctorate at University of California at Davis. He teaches at San Francisco State University.

Acknowledgements

I thank Jonathan Dorsey, Elizabeth Schechter, Kristian Kemtrup, and Christian Coons for helpful philosophical discussions and correspondence. I also thank Brandon Kopald, Brianna Paul, Diana Van Lancker Sidtis, and Sandrine de Ribaupierre for their input regarding surgical, diagnostic, and general neurological matters.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). January, 2016. Volume 3, Issue 4.

Citation

Blackmon, James. 2016. "Hemispherectomies and Independently Conscious Brain Regions." *Journal of Cognition and Neuroethics* 3 (4): 1–26.

Hemispherectomies and Independently Conscious Brain Regions

James Blackmon

Abstract

I argue that if minds supervene on the intrinsic physical properties of things like brains, then typical human brains host many minds at once. Support comes from science-fiction realities that, unlike split-brain cases, have received little direct attention from philosophers. One of these realities is that some patients are functioning (albeit impaired) and phenomenally conscious by all medical and commonsense accounts despite the fact that they have undergone a hemispherectomy: an entire brain hemisphere has been fully detached. Another is the Wada test, in which a patient has each hemisphere anesthetized, one after the other, while the other hemisphere is awake and functioning—again, phenomenally conscious by any standard. I will argue that hemispherectomies, Wada tests, and related procedures each present cases in which the minds that exist after the detachment (or anesthetization) of a hemisphere are surviving minds which must be associated with the surviving (or un-anesthetized) hemisphere. I will argue that such surviving minds existed before the medical procedure, instantiated by the then-intact hemisphere that was due to survive the loss of its complementary hemisphere. If so, then the typical subject has at least three minds: a “left hemisphere mind”, a “right hemisphere mind”, and a “whole brain mind”. But the argument generalizes to cases in which smaller portions of the brain are lost, yielding a great number of additional minds, some overlapping. Some important ethical implications are raised and briefly examined.

Keywords

Hemispherectomy, Wada test, Consciousness, Supervenience, Split-brain

1. Introduction

An *anatomical hemispherectomy* is a medical procedure in which one entire brain hemisphere is surgically removed from the cranium and discarded, leaving the patient with the remaining hemisphere intact and functioning. In a *functional hemispherectomy*, some of a hemisphere is removed while the rest is disconnected and left *in situ*. In a *hemispherotomy*, a hemisphere’s connections to the other hemisphere and to other brain centers are cut, but the hemisphere is left *in situ*. In each of these procedures, one hemisphere is disconnected from the other hemisphere, which remains connected to and functioning in the body. Anatomical and functional hemispherectomies and hemispherotomies are each specific cases of what will here be called *hemisphere*

disconnection, i.e., cases in which one hemisphere has its neural pathways to the other hemisphere and to the body severed.¹

Hemisphere disconnection procedures have been performed on humans since before the 1930s.² They are now used on both children and adults, most commonly as a treatment for severe seizures brought on by unihemispheric damage such as that caused by Rasmussen's Encephalitis, Sturge-Weber Syndrome, and hemimegalencephaly (Bahuleyan et al. 2012). Hemisphere disconnections are fairly radical operations; nevertheless, some hemispherectomy patients, despite the lack of a functioning hemisphere, recover remarkably, going on to complete college, holding regular jobs (Vanlancker-Sidtis 2004), and commenting online on their experiences.

Importantly for our purposes, surgical hemisphere disconnections are distinct from the more familiar "split-brain" phenomenon in which both hemispheres, despite having their connection via the corpus callosum severed, are connected with the rest of the brain as well as with the body via functioning sensory and motor pathways.^{3,4} Split-brain patients have two functioning hemispheres which receive sensory data and send motor commands; hemispherectomy patients do not. Unlike split-brain phenomena, which philosophers have been investigating since Nagel's classic (1971), hemispherectomies and hemispherotomies have received comparatively little attention in philosophy.⁵

The intracarotid amobarbital procedure, commonly named the *Wada Test* after its originator, the Japanese Canadian neurologist Juhn Atsushi Wada, is a medical procedure which successively anesthetizes each hemisphere while the other hemisphere is awake

-
1. See de Ribaupierre and Delalande (2008) for an overview of these various surgical techniques which share as a "common goal" "the interruption of the corpus callosum, the internal capsule and corona radiata, and the mesial temporal structures as well as the frontal horizontal fibers."
 2. Citing Dandy (1928), de Ribaupierre and Delalande (2008) report that anatomic hemispherectomy was first performed in the late 1920s.
 3. For early papers on split-brain research, see Sperry (1964, 1966), Gazzaniga (1967) and Gazzaniga and LeDoux (1978). Gazzaniga (2000) provides a more recent overview.
 4. Following Bayne (2008) we can refer to both the commissurotomy (in which the corpus callosum along with other connections are severed) and the callostomy (in which only the corpus callosum is severed) as "split-brain" procedures.
 5. Although, see Marks (1981), Puccetti (1993), Schechter (2012).

and functioning (Wada 1949, Snyder and Harris 1997).^{6,7} The point is to determine the language and memory capacities of each hemisphere for people who may undergo hemisphere disconnection surgery to treat conditions such as severe epilepsy. In the Wada test, sodium amobarbital or another barbiturate is introduced to one hemisphere via the carotid artery while the other hemisphere is left awake and functioning. At this point the patient performs language and memory tasks, giving normal indications of consciousness albeit along with measurable cognitive impairment, so that clinicians can assess the cognitive capacities of a brain with one hemisphere anesthetized. Once this anesthetized hemisphere awakens, the process is repeated for the other hemisphere. As Snyder and Harris (1997) point out, the Wada test is in effect a “reversible lesion”, and as such it provides additional proof that each hemisphere can cognitively function independently of the other. It is also evidence that such independent functioning can be temporary, starting abruptly and lasting only a brief period of time (often only a matter of minutes), and that it can be something that is consciously experienced by something that does not include the anesthetized hemisphere.⁸ Philosophers appear to have given the Wada test virtually no attention. I will include the Wada test as another specific case of “hemisphere disconnection”, even though in this case the disconnection is not surgical and is only temporary.

The medical literature makes clear that, medical tragedies aside, patients of hemisphere disconnection are conscious beings—at least insofar as one medically detects consciousness in living things.⁹ Granting this much and assuming a fairly standard view about the mind-body relation, I will argue here for the thesis that any whole and functioning brain provides the supervenience base for many phenomenally conscious minds, some overlapping, some not.¹⁰

-
6. Typically, the procedure perfuses only two (anterior and middle) of the three cerebral arteries and so does not normally anesthetize the entire hemisphere, although in some cases this does occur. This fact should be kept in mind when considering the data we get from most actual Wada tests. However, it should also be kept in mind that anesthetization of an entire hemisphere is possible, even if rare.
 7. A similar procedure, though differing in methods and clinical goals, was independently developed by W. James Gardner (Snyder and Harris 1997).
 8. Meador and Loring et al. (1997) employ a modified version of the Glasgow Coma Scale (Teasdale and Jennett 1974) in order to assess levels of consciousness in patients undergoing the Wada test.
 9. The logical or conceptual possibility of philosophical zombies may remain, but if so, it is hardly any more of a problem in the case of patients of hemisphere disconnection than it is in the case of normal humans.
 10. The argument given here is not intended to challenge particular arguments regarding the “Mental Problems

My argument depends upon a principle which, though widespread, remains controversial. This principle is a particular mind-body supervenience thesis (MBS).

MBS: The phenomenal properties instantiated by a physical thing supervene on its intrinsic physical properties.¹¹

MBS is accepted by many theories of mind. The principle is trivially true if, as identity theory holds, mental properties are identical with physical properties.¹² Other views that deny the identity of these properties (emergentism, epiphenomenalism, property dualism, and a variety of forms of functionalism) can and often do accept MBS.¹³ Because I argue here that if MBS true, then each functioning human brain hosts multiple minds, the foregoing reasoning works under the assumption of MBS except where otherwise indicated.

Before presenting the argument, I should say why, even with all the scrutiny that split-brains have received, hemisphere disconnection cases warrant more attention than they have been getting by philosophers. Granted, hemispherectomy cases are closely related to split-brain cases in many ways, so one might suspect that hemispherectomies just provide an additional medical instance of an issue that has been in the philosophical literature since Nagel (1971). For this reason, important differences between hemisphere disconnection cases and split-brain cases should be noted.

First, the cases of hemisphere disconnection permit a more direct interpretation of the research data they provide. For obvious reasons the data we have on hemispherectomy patients are unequivocally data about what the surviving hemisphere, not the removed or detached hemisphere, can do. As Schechter (2012) notes, the hemispheres are not entirely split in split-brain cases, creating etiological ambiguity. For example, the possibility of interhemispheric cortical transfer of information, by way of the superior colliculus, remains in split-brain cases (Savazzi et al. 2007; Roser and Corballis

of the Many" or the "Many Thinkers Problem", although the thesis here obviously rejects the view that more than one mind per brain serves as a *reductio*. See Unger (2004), Kovacs (2010).

11. The supervenience relation I am using is *weak individual* supervenience: A-properties weakly supervene on B-properties if and only if for any possible world *w* and any individuals *x* and *y* in *w*, if *x* and *y* are B-indiscernible in *w*, then they are A-indiscernible in *w*. I take the term 'phenomenal' to rule out mental properties about which one can reasonably be an externalist; however, see Lycan (2001).
12. If A-properties are identical to B-properties, then (trivially) for any possible world *w* and any individuals *x* and *y* in *w*, if *x* and *y* are B-indiscernible in *w*, then they are A-indiscernible in *w*.
13. Granted, there are those who explicitly reject MBS; an externalist about phenomenal states (see Lycan (2001)) would count, and obviously an interactionist dualist would count, as well.

2001). Moreover, the outputs of each hemisphere may be integrated by some non-cortical structure (the cerebellum might coordinate bimanual actions) or by perceptual cues to which both hemispheres have access (Seymour et al. 1994; Ivry et al. 2002). Thus, determining exactly whether and how each hemisphere contributes to the behavior studied is a research problem in split-brain cases. But there is not even the possibility of interhemispheric transfer or non-cortical integration in the case of hemispherectomy; this is simply because there is no other connected hemisphere. Hemispherectomy cases, along with the other forms of disconnection, thus avoid much of the ambiguity that arises in split-brain research.

Second, related to the first point, cases of hemisphere disconnection avoid theoretical ambiguity. Regarding the data on split-brain cases considered in ignorance of hemisphere disconnection cases, it remains plausible that the split-brain patient has only one mind and that no mind could be associated merely with one or the other hemisphere. But in light of hemisphere disconnection cases, maintaining this view requires some reaching. I grant the epistemic possibility that a split-brain patient has one mind while a hemisphere disconnection patient has none, but I also think that this view requires (and will long await) additional empirical support, enough to outweigh the reasons we currently have to think otherwise.

Third, hemisphere disconnection introduces ethical issues which do not arise in split-brain cases. While both hemispheres of the split-brain patient continue to live and interact with the world, this is clearly not so with hemisphere disconnection. Recall that by an anatomical hemispherectomy, a hemisphere is removed from the cranium and discarded, while by a functional hemispherectomy or a hemispherotomy, most or all of a hemisphere is detached but left in the cranium. Thus, if these things can be independently conscious, then it is possible that conscious things are being killed or put into sensory-motor isolation. We will return to the ethical implications of letting an arguably conscious hemisphere or other brain region die or of leaving it isolated in a cranium. For now it should be clear that cases of hemispherectomy have an ethical dimension that split-brain cases do not have.

In light of these ways in which hemisphere disconnections are different from split-brain cases, and in light of the scientific, medical, and ethical ramifications considered here, we should be careful not to generalize too broadly from a philosophical literature focused largely on split-brain research. Much has been written about how split-brain cases bear on our notions of consciousness, minds, personhood, and agency. Not all of this literature is responsive to hemispherectomies, hemispherotomies, and Wada tests.

I should also be clear, before moving on, about what I am not trying to do here. For, because this investigation will look a lot like some more familiar investigations into split-brain cases and because those investigations are often concerned with other issues I will not be directly concerned with, there is the possibility that readers will expect this paper to adjudicate on these issues. So I will state upfront four things I am not trying to do with the arguments presented here.

First, I am not trying to advance a robust theory of what it is to have or be a mind. I use the predicate 'has a mind' here to identify conscious entities with sufficient cognitive capacities, leaving it open exactly what those capacities may be.¹⁴ Admittedly, my thesis reveals that I am willing to think of things that have minds as things that can have proper parts that have minds, and this undoubtedly is a controversial claim resting on a controversial concept of the mind. For, some people hold that it is essential to the concept of a mind that minds cannot have minds as proper parts or that mind-having things cannot consist of mind-having proper parts.¹⁵ However, I do not think much turns on this. If someone were to insist that for metaphysical, conceptual, or linguistic reasons the things I am talking about cannot be designated as "minds", I would propose replacing each instance of the term 'mind' with 'independent conscious (and cognitive) entity' or, if that sounds too much like a substance view, to limit myself to the predicate 'is independently conscious and cognitive' where I would have 'has a mind'. These issues are interesting, but I do not see that they bear on the current issue.

Second, I am not proposing to conceptually analyze personhood or agency. Whether the brain regions I will be talking about are associated with persons or agents is an important question in its own right, one which deserves investigation, but the concepts are matters of extensive controversy, and attempting to adjudicate on them here would introduce distractions and unnecessary burdens to my specific task, which is to argue simply that some brain regions are independently conscious entities.¹⁶ Note that even if S

14. It will become clear that I do not think language capacity is a necessary condition for having a mind as I use the term.

15. A typical response to Block's (1978) "Chinese Nation" thought experiment is that even though the system is functionally equivalent to a brain, it does not have a mind. Perhaps this response is often driven by the intuition that either the whole system has a mind or certain of its parts have minds, but not both. Of course, Block's hypothetical system is composed of people, each of which has a mind, thus, according to this intuition, the entire system cannot have a mind.

16. Regarding new work on the issue of personhood for split-brain cases, see Tye (2003), Bayne (2005), and Schechter (2009).

is not a person or an agent on one or more of the many views of personhood or agency in play, *S* may nevertheless be conscious. We can of course ask whether, for example, dogs are conscious without asking whether they are persons or agents. Moreover, two philosophers can disagree on issues of personhood and agency regarding dogs while agreeing that dogs are conscious, and a third philosopher can take issue with their shared notion that dogs are conscious leading to a new debate among the three that might never mention personhood or agency. I think the same possibilities hold of hemispheres and some other brain regions, and I hope to advance my argument without relying on these concepts. I will eventually evoke the concept of personhood, but only in an opaque sense, simply noting that if a hemisphere is a person (leaving unaddressed whatever it is to be a person), then in addition to the ethical matters raised by the hemisphere's alleged independent consciousness are ethical matters raised by its personhood.

Third, I am not trying to address the issue of whether the entities I describe as conscious are conscious in a way that is unified.¹⁷ Like personhood and agency, the concept of the unity of conscious is ambiguous and controversial. And, like personhood and agency, it is also secondary to my particular concern. After all, a unified consciousness presupposes consciousness. If my argument convinces people that typical human brains have many independently conscious regions, then perhaps we can move on to consider the question of whether those regions are conscious in a way that is unified or not. We can also perhaps move on to consider the question of how those independently conscious regions could compose a brain that allegedly has unified consciousness. But again, I think that this question is beside the present point, which is to determine whether these regions are conscious at all. Granted, it does sound like a threat to the concept of the unity of consciousness to say that each of us has many minds, and perhaps it is. But there is no question of the unity of an entity's consciousness, if there is no consciousness, so I propose to address this more basic question, leaving the rest for future investigation.

Fourth, I am not going to explore the question of independently conscious hemispheres with much of an eye to the asymmetries between the hemispheres. Undoubtedly, the hemispheres each tend to display differences in cognitive specialization (Gazzaniga 2000). However, nothing has shown that one hemisphere is less deserving of the designation *conscious mind*. What it is like to be a left hemisphere may be interestingly different than what it is like to be a right hemisphere, but there is nothing it's like to be a hemisphere unless hemispheres are conscious. So again, my aim is simply to address this more basic question.

17. See Bayne (2008).

2. The Arguments from Hemisphere Disconnection

These arguments all begin with the notion that there is some continuity of conscious experience which survives the disconnection of some region of the brain. The proposition might appear controversial until we consider that we are all in a position to know what it's like to lose at least some parts of the brain.¹⁸ This is simply because we have all experienced phenomenally conscious life as our neurons naturally succumb to programmed cell death. I do not mean that we are necessarily in a position to detect and track the loss of neurons by attending to our conscious experience. For many of us, the experience of losing a few neurons over some period of time may seem no different than the experience of having the neurons we currently have. For others less fortunate—victims of stroke or traumatic brain injury—the differences between what it was like to have had those neurons and what it is like now to lack them is stark. Nevertheless, in both cases, conscious experience can continue.¹⁹

Particularly striking examples of knowing what it's like to lose neurons come from cases of surgical hemisphere disconnection in which the surviving subject is still phenomenally conscious and has the cognitive capacities to reflect on his or her current and prior phenomenal states. Many of these patients, incidentally, are awake and responsive during the procedure, this being the ideal condition for brain surgery of this kind. The argument can now be made regarding such patients.

Sometimes, there is something it's like to lose a brain hemisphere, a “detachment experience”. But this detachment experience requires experiencing interaction with that hemisphere and experiencing subsequent lack of interaction with that hemisphere.²⁰ Experiencing interaction with that hemisphere supervenes on the intrinsic properties of some part of the brain that will survive the surgical disconnection—the other hemisphere or some proper part of it. After all, any part that includes any part of the hemisphere to be lost is not *interacting* with it and is not a thing that will lose it. Thus, the healthy hemisphere due to survive (or some proper part of it) supplies, prior to the disconnection,

18. Throughout, I follow Nagel (1974) in using the phrase *what it's like* to capture the subjective or phenomenal nature of conscious experience.

19. In the same respect, we know what it's like to have a prefrontal cortex metabolizing glucose, even if we do not identify our experiences as such. Moreover, the phenomenal experiences of headaches, thumb aches, tinnitus, and *déjà vu* (to name only a few) each constitute what it's like to have a brain in a particular kind of physiological state even if we are ignorant about the neuroscience of those states.

20. Of course, as with a headache or tinnitus and the like, the experience doesn't have to be identified as corresponding to any neural phenomena.

the supervenience base of the experience of interacting with the other hemisphere. And therefore, the healthy hemisphere due to survive (or some proper part of it) has experience prior to the disconnection. A specific version of this general argument follows.

The Argument from Hemispherectomy

1. In some cases, there is something it's like to have a brain hemisphere detached, a detachment experience.
2. A detachment experience consists in an interaction experience and a subsequent lack of interaction experience.
3. This interaction experience supervenes on the intrinsic properties of some part of the brain that will survive the hemispherectomy—the other hemisphere or some proper part of it.
4. Thus, that other hemisphere or some proper part of it has experience before (and after) the hemispherectomy.

From this we can now conclude that a healthy typical brain gives rise to at least two additional and different phenomenally conscious minds: one associated with the left hemisphere, another associated with the right. The mind that would survive the detachment of the right hemisphere is the mind associated with the left hemisphere (or some part of it). The mind that would survive the detachment of the left is the mind associated with the right hemisphere (or some part of it).

One more specific argument is worth presenting in explicit form. Recall that in the Wada test, brain hemispheres are independently anesthetized so that clinicians can diagnose some of the cognitive capacities of the other hemisphere, which remains awake and functioning. Patients now post reports of their experiences online, commenting on what it was like, for instance, to attempt to name objects with one or the other hemisphere anesthetized.²¹

21. Current sources of patient reports can be found at sites such as www.epilepsy.com. Patients have also begun posting videos online in which they recount their experience.

The Argument from Wada Test

1. In some cases, there is something it's like to have a brain hemisphere anesthetized while the other hemisphere is not anesthetized.
2. Experiencing the anesthetization of a brain hemisphere requires having experienced interaction with that hemisphere and experiencing subsequent lack of interaction with that hemisphere.
3. Experiencing interaction with that hemisphere supervenes on the intrinsic properties of some part of the brain that will survive the hemispherectomy—the other hemisphere or some proper part of it.
4. Thus, that other hemisphere or some proper part of it has experience before (and after) the anesthetization.

In addition to the Argument from Hemispherectomy and the Argument from Wada Test, there is an analogous Argument from Minor Stroke, in which the premises involve what it's like to lose interaction with a small collection of neurons due to a minor stroke. Consider each region of your brain that you might lose to some stroke from which a conscious survivor would emerge. On the current view, for each such region, there is the complementary brain region associated with an independently conscious mind, many of these minds not yet aware of what it would be like to lose the region with which they are so intimately tied. There is also an analogous Argument from Cell Death, in which the premises involve what it's like to lose interaction with a perhaps small and scattered collection of neurons due to programmed cell death or some other cause. In one way, these arguments, which involve phenomena that are less extreme and more familiar, might be more convincing than those regarding hemispherectomies and Wada tests; after all, no one worries that the surviving and functioning brain, having shed a few neurons due to a minor stroke, programmed cell death, or the like, is not conscious.

Trivially, a generalization of these arguments concludes that there are far more minds—many billions, as absurd or unnerving as one might find that conclusion to be.²² Enumerating neurons, there is the mind instantiated by the entire brain-minus-neuron-1,

22. Unger (2004) expresses such a reaction. This is addressed shortly.

the one instantiated by the entire brain-minus-neuron-2, and so on. And we can do this for the brain minus certain neuron pairs, neuron triplets, and so on up to whatever limits would preclude granting minds to something we might consider to have “insufficient brain matter or structure” for the instantiation of a conscious mind. Without worrying here about how we would draw the line between sufficient and insufficient brain matter or structure, we can at least extend the argument to the conclusion that each normal brain gives rise through its many proper parts to many minds.²³

The view that each normal brain gives rise through its many proper parts to many minds will be used by some in a *reductio*. For example, Nagel (1971), in considering the possibilities raised by split-brain cases, writes:

In case anyone is inclined to embrace the conclusion that we all have two minds, let me suggest that the trouble will not end there. For the mental operations of a single hemisphere, such as vision, hearing, speech, writing, verbal comprehension, etc. can to a great extent be separated from one another by suitable cortical disconnections; why then should we not regard each hemisphere as inhabited by several cooperating minds with specialized capacities? Where is one to stop? If the decision on the number of minds associated with a brain is largely arbitrary, the original point of the question has disappeared.

Nagel is concerned that if we accept the existence of more than one mind per typical human brain, then the number of minds we acknowledge is arbitrary. But this does not follow. What may follow from accepting the existence of more than one mind per typical human brain is that we do not yet know what counts as sufficient brain matter or structure for the realization of a mind. But this is nothing new. We already knew that we do not yet know what counts as sufficient brain matter or structure for the realization of a mind.

Unger (2004) finds the proposal that there are numerous minds or numerous “experiencings” associated with what we normally take to be one human individual to be “incredible”, “absurd”, and “disturbing”. He holds that “there’s nobody, I trust, who thinks there are many billions of experiencings physically promoted largely by the left hemisphere, and billions more largely promoted by the right,” admitting that he doesn’t

23. It is reasonable to suppose that the line should be drawn somewhere. Otherwise, one would have to defend the view that a single neuron instantiates a “mind” which could experience being detached from the rest of the brain. Not even a proponent of panpsychism has to embrace this conclusion.

think the view is “all that plausible”. According to the arguments considered here, however, there are in deed many billions of “experiencings” promoted by parts of the brain.

The following objections will clarify the premises, hopefully preventing some objections that might be based on a misunderstanding of the argument.

2.1 Premise 1: The Existence of Such Experiences

How do we know there is something it is like to lose an entire brain hemisphere? After all, losing an entire brain hemisphere is drastically different, from a purely physiological point of view, than losing some neurons. In the latter case, we are left with what still counts (by all standards, medical, scientific, and common sense) as an entire functioning brain, while in the former, we are not. Perhaps then, there is nothing it is like to lose an entire brain hemisphere, just as there would be nothing it would be like to become a philosophical zombie, or much as there might be nothing it is like to lapse into unconsciousness, or to die.

Admittedly, in the strict spirit of philosophical skepticism, most of us do not know that there is something it is like to lose an entire brain hemisphere simply because most of us have not undergone a hemisphere disconnection and survived to contemplate the results. So, the first premise can be doubted by most of us, if we like. However, to doubt the first premise is to invoke a special application of the problem of other minds, an application which is perhaps a bit further justified by the fact that hemisphere disconnection patients exhibit diminished cognitive capacities. So, as with the standard problem of other minds, we have recourse to analogy from our personal experience to the experiences of others. And as previously established, we all know what it’s like to lose some neurons. Many of us also know what it’s like to have significantly diminished cognitive capacities due to exhaustion, illness, injury, medication, or age. Thus (or so I would argue), we have at least some way of gauging by analogy what it is like to move along the relevant physiological spectrum of possible states for a brain or any composite of functioning neurons.

But more importantly, for the purposes of our argument, we do not need to know what it is like to survive a hemisphere disconnection; we just need to know that it’s like something. Given that some hemispherectomy patients are conscious by any operative standard, hemisphere disconnections do not present a special kind of problem.

Importantly, although Nagel’s (1971) paper is almost entirely limited to split-brain phenomena, he does briefly consider the possibility of a person deprived of the

left hemisphere, which at the time was believed by many to be essential to language processing.²⁴ His point is that even if a person had only the hemisphere which does not predominantly process speech, there is no reason to deny that the person is conscious.

There is no doubt that if a person were deprived of his left hemisphere entirely, so that the only capacities remaining to him were those of the right, we should not on that account say that he had been converted into an automaton. Though speechless, he would remain conscious and active, with a diminished visual field and partial paralysis on the right side from which he would eventually recover to some extent. In view of this, it would seem arbitrary to deny that the activities of the right hemisphere are conscious, just because they occur side by side with those of the left hemisphere, about whose consciousness there is no question.

Nagel's point stands today. While the capacity of a hemisphere for coherent, interactive speech may remove any empirical doubt about its consciousness, each hemisphere alone has capacities sufficient for the justified attribution of consciousness.

2.2. Premise 2: Experiencing Detachment of a Hemisphere

First, the argument treats the experience of detachment from *x* as a composite experience that must include some experience of being attached to *x* and some subsequent experience of not being attached to *x*. Such a composite experience obviously spans time. The argument does not require that the very moment of detachment itself be consciously experienced or identified as such.²⁵ In the same way, one might experience the loss of one's wallet simply by first experiencing the having of the wallet, then experiencing the lack of the wallet. The very moment of loss does not need to be identified as such in that moment or even afterward; one just realizes that at some point the wallet was lost, and this requires having experienced its presence followed by an experience of its absence.

24. We now know that language dominance is only typically lateralized to the left hemisphere. While left-hemisphere language dominance is estimated to exist in well over 90% of right-handers, Knecht (2000) reports that left-hemisphere language dominance can reach up to 27%.

25. As previously noted, the procedure typically involves keeping the patient awake and responsive throughout, which is considered to be ideal because it allows surgeons to track the cognitive and functional results of the operation as it progresses.

Second, experiencing the detachment from x , as construed here, also involves not something's "losing a proper mereological part of itself", but something's detaching from that complement, x , which was once attached and interacting and later is not.²⁶

Finally, my argument is not about entities that believe they have experienced loss, but about things that truly have experienced loss even if they have not acknowledged it as such. We may grant that we cannot truly know at any moment whether we have actually experienced loss. Confabulation, amnesia, and other forms of faulty memory may lead me to believe that I have experienced something that I have not experienced or that I have not experienced something I have experienced.²⁷ One's believing that one has experienced loss and one's experiencing loss are different things. Consider, for instance, the experience of coming to believe that you have lost some region of the brain, and suppose that, despite your belief, you never had the brain region in the first place. This is not a detachment experience, and as such it is not addressed by the argument. The present argument applies to entities that truly have had conscious experience of loss.

2.3 Premise 3: Supervenience and Causal Interaction

The phenomenal experience of interacting with some x supervenes on the intrinsic physical properties of something else that does not overlap x .²⁸ Brain hemispheres conjoined by a functioning corpus callosum causally interact with each other. According to this premise, each brain hemisphere, not the whole brain, instantiates the phenomenal experience of interacting with the other hemisphere.

Churchland (1981) muses about the possibility of linking human brains to each other much as hemispheres are linked to each other via the corpus callosum: "Once the channel is opened between two or more people, they can learn (*learn*) to exchange information and coordinate their behavior with the same intimacy and virtuosity displayed by your

26. One might invoke mereological essentialism in order to defend this view; however, that is not necessary. We need only accept that there is some y which once interacted (through physical attachment) with x and now does not.

27. As Russell famously point out, for all we know, we (along with everything else in the universe) have only just now popped into existence with false memories of having lost something which in fact we never had. In that case, we think we have lost something, but we haven't lost it.

28. Of course, the intrinsic properties of x help to determine the nature of the interaction and thereby can influence which phenomenal experiences are had by the thing that is interacting with x , but these properties of x are effective only insofar as they change the intrinsic physical properties of the thing that is interacting with x .

own cerebral hemispheres. Think what this might do for hockey teams, and ballet companies, and research teams!” Churchland soon after asks, “How will such people understand and conceive of other individuals?” His answer: “In roughly the same fashion that your right hemisphere ‘understands’ and ‘conceives of’ your left hemisphere—intimately and efficiently, but not propositionally!” For our purposes, we might ask what it would be like to become linked to another brain via a connection as complex as the corpus callosum. Perhaps we currently have no justifiable specific answer. Even so, we can at least grant that it would be like something.²⁹ Perhaps, then, whatever it would be like to be a brain linked to another brain is much like what it is like to be a hemisphere linked to another hemisphere. And even if not, we can acknowledge that just as there should be something it is like to be a brain linked to another brain, there should be something it is like to be a hemisphere linked to another hemisphere.

None of this should be taken to mean that a human with both hemispheres connected by a healthy corpus callosum can report on what it is like for one hemisphere to interact with the other. Recall that all of us have lost neurons. This, of course, does not mean that prior to the loss we can report on what it is like for the overwhelming majority of the brain to interact with some small and scattered portion soon to be lost. Similarly, a stroke patient, prior to the stroke, cannot be expected to be able to make such a report. Reports are the product of complex causal relations holding among complex neural entities. Any reports an integrated hemisphere might be disposed to make about what it’s like to interact with the other hemisphere are conceivably subject to being “washed out” by or combined with other activity, including any other reports that the other hemisphere would be disposed to make. We have good evidence for this. Differences in the dispositions of each hemisphere to report are in fact just what the split-brain research so famously shows.³⁰ And fortunately for our purposes, the Wada test provides additional evidence about what it’s like for one hemisphere to experience disconnection from the

29. For, the alternative, that it is like nothing, requires that we implausibly decide that brains lose consciousness upon becoming linked to other brains.

30. In a paradigm example (Scientific American 2002), “Joe”, a split-brain patient has the word ‘toad’ very briefly presented on his left visual field (processed by the right hemisphere) and ‘stool’ very briefly presented on his right visual field (processed by the left hemisphere). When Joe is asked to draw what he saw with his left hand (controlled by the right hemisphere), he draws something resembling a toad—certainly not a stool nor a toadstool. When asked to say what he saw, he says, “Stool” (this speech act processed by the left hemisphere). At this point, Joe’s left hand (controlled by the right hemisphere which has now heard the left hemisphere’s production of the word ‘stool’) draws a three-legged stool. Joe does not draw a toadstool.

other, and *vice versa*, this being largely the point of the Wada test. And unlike split-brain research, the Wada test also provides evidence about what it's like for one hemisphere to experience reintegration with the other hemisphere as it awakens, and *vice versa*. Thus we do have a promising window into what it's like to interact with the other hemisphere, for the Wada test allows one hemisphere to independently supply reports and behavior indicating what it can and can no longer see, identify, control, and remember. The deficits revealed presumably testify as to what the still-functioning hemisphere is accustomed to, even if it never independently considered what it was accustomed to prior to the Wada test.

3. Objections

Any argument that concludes that brains have an indeterminate and large number of minds associated with it, or that regions of the brain are independently conscious, is assured to encounter objections. Earlier, the hesitations of Nagel and the strong sense of bewilderment expressed by Unger were mentioned. The sense that this view is implausible or even absurd cannot easily be dismissed. For most of us, daily experience does not seem to involve a cacophony of minds, and that very fact appears to serve as a sufficient objection to the view. While I suspect that this kind of objection begs the question in simply assuming that there is a special single conscious mind which would somehow experience competing others if there were any, we will not examine it here. I accept the strange feel of the thesis that brains have many minds, but to the extent that this is an objection, it is so far only an objection to the conclusion of my argument. It does not tackle any step of reasoning along the way.

Other objections do however address the general argument presented here. We will consider three. Each of them is an objection to Premise 3, the claim that interaction experiences supervene on the intrinsic properties of some part of the brain that will (or would) survive a disconnection, that is, the other hemisphere or some proper part of it.

3.1 Objection from Considerations of Personhood

One might object that the functionally connected hemisphere is not independently associated with its own person, and because of this, it has no conscious experience. This objection to Premise 3 requires some nontrivial philosophical commitments to what it is to be a person, and among them must be a commitment to the proposition that a person cannot be associated with a brain hemisphere if it is functionally connected to

another brain hemisphere.³¹ Clearly, the proposition itself is not *prima facie* part of our normal conception of a person, at least not in the sense that standard attempts to define personhood make any mention of brain hemispheres or their connectedness. Ideally then, the proposition would be compelled by something else in the concept of personhood.

While I reject the view that consciousness requires personhood (for the reasons previously indicated), I will try to make the most of this objection by considering what appears to be a reasonable version of it. Plausibly, the objection might include the proposition that persons cannot overlap or that one person cannot be a proper part of another person. After all, we do not typically think of persons as entities that can truly overlap in the physical sense considered here.³² So perhaps there is some promise for the view that a connected hemisphere is not a person. I have two responses to this version of the objection.

First, even if a functionally connected hemisphere is not a person, this does not entail that it is not independently conscious. As previously noted, it is quite plausible that some animals are independently conscious without being persons according to any extant conception of personhood. The objection, then, puts the cart before the horse: Personhood requires at least the capacity for experience; the converse does not obviously hold. Thus we can concede this much of the objection—that a functionally connected hemisphere is not a person—and nevertheless retain the possibility of consciousness for such hemispheres.

Second, the thesis that a functionally connected hemisphere cannot be associated with a distinct person results in a dilemma: Either a hemisphere, functionally connected or not, cannot be associated with a distinct person, or a hemisphere, when functionally

31. Though see Schechter (2012).

32. Someone might want to champion alleged counterexamples to the proposition that persons cannot overlap: the pregnant woman, the conjoined twins, the “split personality”, the soulmates... While this would not be the place for an extended argument, I will say that I do not think that, on closer inspection, any of these cases would count as counterexamples. Even in the normal cases of conjoined twins, while we must acknowledge that can they share organs, we typically think of them as two individual persons who are biologically connected, even dependent, on each other. However, there do seem to be reasons for rejecting the proposition that persons cannot overlap. Schechter (2012) makes a much more compelling case for overlapping persons associated with the brains and hemispheres of split-brain patients. There is also the possibility that conjoined twins share a brain region, as with the craniopagus conjoined twins Krista and Tatiana Hogan, whose brains are believed to be connected by a “thalamic bridge” and whose behavior suggests to some that they share sensory experience (Dominus 2011). Both Schechter’s considerations and the anecdotal and anatomical evidence of Krista and Tatiana Hogen present stronger counters to the proposition that persons cannot overlap.

connected, cannot be associated with a distinct person. On the first branch, hemisphere disconnection patients are not persons. But our research and the behavior and reports of patients themselves clearly indicate otherwise. On the second branch, a hemisphere disconnection patient can be a person; however, should that hemisphere ever be functionally connected with another hemisphere, that original person associated with that hemisphere is *ex hypothesi* annihilated. Recall that the Wada test, in which brain hemispheres are independently anesthetized, one after the other, is a medical reality. Thus on the second branch of the dilemma, it would appear that various persons would have to come and go as the test progresses. The reader is left to weigh these options against the proposition that brain hemispheres have independently conscious minds.

3.2 Objection from Interactionist Substance Dualism

According to interactionist substance dualism, the mind is something that can be one way or another without any (immediate) differences in the brain but which interacts somehow with the brain. On this view, it may be the case that only single minds interact causally with brains or brain portions so long as the portion is functioning well enough. Perhaps then cases of hemisphere disconnection are cases in which there is only one mind, one which once interacted with a whole brain but now interacts only with a hemisphere.

This objection also rejects Premise 3. But it does so only by rejecting MBS, and as such, it is strictly not an objection to my view that under the assumption of MBS we have good reason to hold that typical brains host many minds at once. However, a response to this objection does not need to be merely dialectical; two further points can be made.

First, interactionism is *prima facie* an empirically testable claim. After all, showing that a brain or hemisphere exhibits physical effects that cannot be accounted for by any of the extant physical causes would count at least as evidence that the brain was affected by some nonphysical interaction. The brain, on this view, is like a remote-control toy, that toy being something that could be discovered to be under remote control even by scientists who had no conception of electromagnetism or any ability to directly detect it. Thus, unless one holds an empirically untestable form of interactionism, the view embraced by this objection can be placed among the scientific hypotheses awaiting empirical testing.

Second, interactionist substance dualism raises additional questions. One must eventually address the question of what would happen in a case in which two healthy hemispheres are entirely disconnected from each other. Does the single immaterial mind once associated with the whole brain now interact with one, the other, both, or neither?

Again, unless the alleged interactionism is untestable, we should expect some intriguing results. And if it is untestable, we should expect some independent reasoning to favor one option over the other, noting that three of these options entails the existences of “zombie hemispheres”. Admittedly, I have not ruled out any of the options, but do hope to have shown where the burden lies.

3.3 Objection from Shared Anatomical Parts

Hemisphere disconnections leave shared brain structures, such as the brain stem, in place. Perhaps then the supervenience base of the mind that exists after the disconnection is not the properties of the surviving hemisphere, but of the brain stem. If so, then one might argue that a single mind is associated with one of these shared anatomical parts, and the loss of a hemisphere is just the loss of some peripheral, though elaborate, organ for processing information. Note that this objection does not simply argue that a mind can be associated with some non-cortical brain structure; it must further argue that no mind is associated with a hemisphere. Once again, Premise 3 is allegedly false. Like the appeal to interactionist substance dualism, this objection offers another way out. Unlike the appeal to interactionist substance dualism, it appeals only to the physical world. However, it faces two problems.

First, it is not clear that any empirical research supports the idea that our brain hemispheres are not conscious. According to this objection, there is nothing it is like to be a single hemisphere or both hemispheres, and consequently there is nothing it would be like to be a hemisphere or pair of hemispheres losing interaction with the brainstem. There are only conscious experiences that involve what it is like (for the brainstem) to be interacting with the hemispheres, and what it is like not to be interacting with them. This is a radical departure from our medical and neuroscientific understanding of the brain. Note how much cognitive processing would be unconscious on this view. For instance, all the executive control and inhibition performed by the prefrontal cortex would be entirely unconscious. This view might conveniently avoid some of the ethical considerations we will look at in the next and final section, for on this view, hemispheres are not themselves conscious; however, the idea that hemispheres are not conscious appears to be scientifically unfounded.

Second, even if it were true that conscious minds are associated not with any part of the cortex, but with some structure which survives disconnections of either or even both hemispheres, this only means that hemispheres (and their parts) are not independently conscious; it does not mean that there are not many minds associated with the brain

or that other proper parts of the brain are not independently conscious. Recall that the arguments about disconnection, along with those involving stroke and cell death, generalize. Thus, even if a subcortical structure, or some special part of it, is the true “seat of the mind”, the possibility remains that that very thing can be divided in uncounted ways and that the main conclusion still holds: The brain hosts many minds.³³

4. Ethical Considerations

The thesis that brains have many minds has fascinating implications for the nature of consciousness and personal identity. Right now, you are reading this sentence. However, at the very same time a “part of you” is experiencing what it is like to be conjoined and interacting with another part which together comprise the entity that is reading the sentence. In fact, there are many such parts. Some of them understand their role as part of the whole sentence-reading entity; others clearly do not, and yet they are conscious entities nonetheless. For they are parts which would survive as conscious experiencers of the loss of other parts, were those other parts to die or become surgically disconnected or anesthetized.

The thesis that brains have many minds also has implications for how we should treat other conscious entities in general. These implications are identified and briefly elaborated here.

As I have argued, each healthy hemisphere instantiates a conscious entity. If so, then the death of a healthy hemisphere instantiates the death of a conscious entity. Of course, actual medical hemisphere disconnections are performed only when doctors have decided that the hemisphere to be detached is significantly deleterious to the rest of the brain. One can hope that actual medical hemisphere disconnections involve only the detachment of hemispheres so unhealthy that there is not much of a conscious entity there to begin with. Nevertheless, we do acknowledge the possibility or actuality of consciousness even in those whose brains are significantly impaired. The question remains as to whether an impaired hemisphere instantiates a conscious entity, and if so, whether that conscious entity is a mind, a person, a potential person, or something else.³⁴ However it is that we might come to understand the status of such hemispheres,

33. One might try to avoid this second point by insisting that the shared anatomical part is such that it loses its associated mind upon losing any single physical part.

34. If one is inclined to reject out of hand the idea that this impaired hemisphere is conscious, one should consider how we should think of a patient who once had one healthy hemisphere and one significantly impaired hemisphere but, due to accidental traumatic hemispherectomy, lost the healthy hemisphere and

it should be clear that the decision to perform or to undergo a hemisphere disconnection is not necessarily just a decision to surrender some cognitive and physical capacities; it may instead be a decision that involves the death or extinction of at least one robust conscious entity, one we would recognize in other contexts as a real human person just as we recognize individuals with significantly impaired brains (hemispherectomy patients included) as real human persons.³⁵

Furthermore, recall that not all hemisphere disconnections involve the removal and death of the detached hemisphere. In the cases of functional hemispherectomy and hemispherotomy, the hemisphere, or large parts of it, remains alive and in the cranium, but it is physically cut off from the other hemisphere, unable to transmit motor signals or to communicate in any direct way. If that brain hemisphere instantiates a conscious entity, however disabled it might be due to whatever condition was deemed to warrant the procedure, then this form of hemisphere disconnection amounts to putting a conscious entity into a perpetual “locked in” state. The experiencing subject merely rides along with the body (and the controlling other hemisphere) as a passenger. If sensory pathways are also cut off, then the experiencing subject is in a perpetual state of sensory deprivation, as well, consigned to live on without any ability to experience the world or take action in it.³⁶ Thus, the decision to perform or undergo such a hemisphere disconnection is possibly a decision to put an experiencing subject into a medically induced permanent state of solitary confinement.

By some lights, these considerations may seem too speculative or even too dark to take seriously. But they are actualities we already knowingly face in other contexts. Some serious cognitive impairments are now treated as if they leave room for yet an experiencing subject. Cases of “covert consciousness” such as locked-in syndrome and anesthesia awareness are now known medical realities (Stins, J. F. and Laureys 2009;

continued to live with only the significantly impaired hemisphere.

35. Note that, when someone decides to undergo the procedure, the hemisphere due to be removed possibly takes part in this decision procedure, opening up the conceptual possibility of “partial self-sacrifice”.
36. Conceivably, a mind associated with a single hemisphere in such a state drifts into unconsciousness or some kind of torpor in which at least there is nothing we could count as suffering being experienced. If so, then functional hemispherectomies and hemispherectomies would be much like anatomical hemispherectomies. One might be tempted to suppose that this is so just for the psychological convenience of not having to worry about potentially isolated experiencers. But this would clearly be unjustified both rationally and ethically.

Cruse, D. et al. 2011). We must take seriously the epistemic possibility that some impaired hemispheres are conscious.³⁷

Meanwhile, although the question of when it is morally permissible to sacrifice one experiencing subject for the sake of another has long been a central focus in ethics, hemisphere disconnections appear to introduce a new dimension to this old issue. For, unless it can be shown that only unconscious hemispheres are ever detached, we now have reason to believe that a hemisphere disconnection entails not merely a risk to the life of the patient and some assured loss of cognitive and physical capacities, but also, for all we know, either the death or the sensory-motor isolation of an experiencing subject (many such subjects) even in cases where the procedure is fully successful.

In broaching these ethical concerns, my aim is not to raise alarms about hemisphere disconnection procedures in themselves, which by all accounts have the obvious potential to greatly improve the lives of the surviving experiencing subjects. But I do think the preceding considerations raise the question of whether they affect other experiencing subjects who (or which) may also have moral standing. If so, it is important that we understand and heed the conditions under which hemisphere disconnections are morally permissible. This is, of course, an important question in its own right. Moreover, addressing this questions should help us shed new light on the old question of when it is morally permissible to sacrifice some conscious entities for the sake of others.

37. Basl (2013) has us imagine the case of a new device which reveals that some patients once thought to be in a persistent vegetative state are actually conscious. Considering the refusal to use this device on others thought to be in persistent vegetative states, Basl writes, "To do so would be to commit an inexcusable wrong, it would be to knowingly ignore a possible person, someone with the same moral status as ourselves. It would be as if we turned our back on a possible person who might be suffering greatly right in front of us when we had the ready means to alleviate that suffering."

References

- Bahuleyan, Biji, and Shenandoah Robinson, Ajith Rajappan Nair, Jyothish L. Sivanandapanicker, Alan R. Cohen. 2012. "Anatomic Hemispherectomy: Historical Perspective." *World Neurosurgery* 80 (3-4): 396–8.
- Basl, John. 2013. "What to Do About Artificial Consciousness." In *Ethics and Emerging Technology*, edited by R. Sandler, 380–392. New York: Palgrave MacMillan.
- Bayne, Timothy. 2005. "Divided brains and unified phenomenology: A review essay on Michael Tye's *Consciousness and Persons*." *Philosophical Psychology* 18 (4): 495–512.
- Bayne, Timothy. 2008. "The Unity of Consciousness and the Split-Brain Syndrome." *The Journal of Philosophy* 105 (6): 277–300.
- Block, Ned. 1978. "Troubles with Functionalism." *Perception and Cognition: Issues in the Foundations of Psychology*, edited by C. W. Savage, 261–325. Minneapolis: University of Minnesota Press.
- Brotis, A. G. 2015. "Hemispherectomy: Indications, Surgical Techniques, Complications, and Outcome." *J Neurol Neurophysiol* 6: 300.
- Churchland, Paul. 1981. "Eliminative Materialism and the Propositional Attitudes." *The Journal of Philosophy* 78 (2): 67–90.
- Cruse, D., and S. Chennu, C. Chatelle, et al. 2011. "Bedside Detection of Awareness in the Vegetative State: A Cohort Study." *The Lancet* 378 (9809): 2088–2094.
- Dandy, W. E. 1928. "Removal of right cerebral hemisphere for certain tumors with hemiplegia." *Journal of the American Medical Association* 90: 823–825.
- de Ribaupierre, Sandrine and Olivier Delalande. 2008. "Hemispherotomy and other disconnective techniques." *Neurosurgical Focus* 25 (3): E14.
- Dominus, S. 2011. "Could Conjoined Twins Share a Mind?" 25 May. New York: *New York Times*.
- Eccles, J. C., 1966. *Brain and Conscious Experience*. New York: Springer-Verlag.
- Gazzaniga, Michael, and Joseph LeDoux. 1978. *The Integrated Mind*. New York: Plenum Press.
- Gazzaniga, Michael. 1967. "The Split Brain in Man." *Scientific American* 217 (2): 24–29.
- Gazzaniga, Michael. 2000. "Cerebral specialization and interhemispheric communication: does the corpus callosum enable the human condition?" *Brain* 123 (7): 1293–1326.

- Ivry, R., R. Spencer, H. Zelaznik, and J. Diedrichsen. 2002. "The cerebellum and event timing." *Annual New York Academy of Science* 978: 302–317.
- Knecht, S., B. Dräger, M. Deppe, et al. 2000. "Handedness and hemispheric language dominance in healthy humans." *Brain* 123 (12): 2512–2518.
- Kovacs, D. 2010. "Is There a Conservative Solution to the Many Thinkers Problem?" *Ratio* 23 (3): 275–290.
- Lycan, William G. 2001. "The Case for Phenomenal Externalism." In *Philosophical Perspectives, Vol. 15: Metaphysics*, edited by J. E. Tomberlin, 17–36. Atascadero: Ridgeview Publishing.
- Marks, Charles E. 1981. *Commissurotomy, Consciousness and Unity of Mind*. Cambridge: MIT Press.
- Nagel, T. 1971. "Brain Bisection and the Unity of Consciousness." *Synthese* 22: 396–413.
- Nagel, T. 1974. "What Is It Like to Be a Bat?" *The Philosophical Review* 83 (4): 435–450.
- Puccetti, Ronald. 1973. "Brain bisection and personal identity." *British Journal for the Philosophy of Science* 24: 339–355.
- Puccetti, Ronald. 1981. "The case for mental duality: evidence from split brain data and other considerations." *The Behavioral and Brain Sciences* 4: 93–123.
- Puccetti, Roland. 1993. "Dennett on the split-brain." *Psycoloquy* 4 (52).
- Roser, M. and M. Cordballis. 2001. "Interhemispheric neutral summation in the split brain with symmetrical and asymmetrical displays." *Neurophysiologia* 40: 1300–12.
- Schechter, Elizabeth. 2009. "Persons and psychological frameworks: A critique of Tye." *Journal of Consciousness Studies* 16 (2-3): 141-163
- Schechter, Elizabeth. 2012. "Intentions and Unified Agency: Insights from the Split-Brain Phenomenon." *Mind & Language* 27 (5): 570–594.
- Scientific American. 2002. "Severed Corpus Callosum." (VT 9501) #7: *Scientific American Frontiers Introductory Psychology Video Collection*, 2nd Edition. New York: Worth Publishers. DVD.
- Seymour, S. and P. Reuter-Lorenz, and M. Gazzaniga. 1994. "The disconnection syndrome: basic findings reaffirmed." *Brain* 117: 105–115.
- Snyder, Peter Jeffrey and Lauren Julius Harris. 1997. "The Intracarotid Amobarbital Procedure: An Historical Perspective" *Brain and Cognition* 33: 18–32.
- Sperry, R. W. 1964. "The great cerebral commissure." *Scientific American* 210: 42–52.

- Sperry, R. W. 1966. "Brain Bisection and Mechanisms of Consciousness." In *Brain and Conscious Experience*, edited by Eccles, J. C., 298–313. New York: Springer-Verlag.
- Stins, J. F., and S. Laureys. 2009. "Thought translation, tennis and Turing tests in the vegetative state." *Phenomenology and the Cognitive Sciences* 8 (3): 361–370.
- Taylor, Jill Bolte. 2008. *My Stroke of Insight*. New York: Viking.
- Teasdale, G., and B. Jennett. 1974. "Assessment of coma and impairment consciousness. A practical scale." *Lancet* (2): 81–84.
- Tye, M. 2003. *Consciousness and persons*. Cambridge, MA: MIT Press.
- Unger, Peter. 2004. "The Mental Problems of the Many." In *Oxford Studies in Metaphysics*, Vol. 1, edited by D. Zimmerman. Oxford: Clarendon Press.
- VanLancker-Sidtis, D. 2004. "When only the right hemisphere is left: Studies in language and communication." *Brain & Language* 91 (2): 199–211.
- Wada, J. 1949. "A new method for the determination of the side of cerebral speech dominance. A preliminary report of the intra-carotid injection of sodium amytal in man." *Igaku to Seibutsugaki* 14: 221–222.

Journal of Cognition and Neuroethics

The Complexity of Suicide: Review of Recent Neuroscientific Evidence

Erica Ching

University of Toronto

Author Note

Erica Ching, Department of Psychology, University of Toronto, 100 St. George Street, Toronto, Ontario, Canada M5S 3G3. This research was conducted under the supervision of Dr. Jonathan Downar, MD, PhD, FRCPC, MRI-Guided rTMS Clinic, Department of Psychiatry, University Health Network; Assistant Professor, Department of Psychiatry, University of Toronto. Correspondence concerning this article should be addressed to Erica Ching, erica.ching@mail.utoronto.ca.

Biography

Erica received her BSc degree in psychology from the University of Toronto, in Toronto, Ontario, Canada. She has been involved in research efforts at both the Centre for Addiction and Mental Health, as well as at Toronto Western Hospital. Her research interests are primarily in treatment-resistant depression, neuroimaging, repetitive transcranial magnetic stimulation (rTMS), and the role of glutamate in mood disorders.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). January, 2016. Volume 3, Issue 4.

Citation

Ching, Erica. 2016. "The Complexity of Suicide: Review of Recent Neuroscientific Evidence." *Journal of Cognition and Neuroethics* 3 (4): 27–40.

The Complexity of Suicide: Review of Recent Neuroscientific Evidence

Erica Ching

Abstract

Two important aspects of human behaviour that become maladaptive in those individuals who are at high risk for suicide include: (a) the ability to engage in self-referential thought; and (b) the ability to feel complex and abstract emotions. Research suggests that the aberrant connectivity between and within important neural structures plays a major role in the problematically altered perceptions often held by highly depressed individuals, which ultimately contributes a great deal to the expression of the symptoms characteristic of the disorder. The present work provides a general overview of findings gathered from recent neuroscientific literature, synthesizing a comprehensive view that emphasizes the importance of interconnected networks in the brain in major depressive disorder (MDD), made possible by recent advances in neuroimaging technology. Of particular importance appears to be the ventromedial prefrontal cortex (VMPFC), the dorsomedial prefrontal cortex (DMPFC), and the pre/subgenual anterior cingulate cortices (pgACC/sgACC), the amygdala, and the insula. It is suggested that, rather than being basic diagnoses that can be identified through confirming items on an inventory of symptoms, major depressive disorder may be viewed as a malfunction in an intricate system of networks in the brain.

Keywords

Suicide, Depression, Neuroimaging, Prefrontal Cortex, Anterior Cingulate Cortex, Amygdala, Insula, Networks

The Complexity of Suicide: Review of Recent Neuroscientific Evidence

In the last two decades, convergent research on suicidal behaviour has made use of core concepts from a variety of different theoretical models, including the diathesis-stress (Mann et al. 1999), social interactionist (Lam et al. 2010), cognitive (Dieserud et al. 2001), and molecular models (Mann 2003; Turecki 2014). However, there has been little done toward integrating these diverse models of behaviour to synthesize a comprehensive neuropsychological view, despite the recent advances in neuroimaging technologies that have been made. Many neuroimaging studies focus on neurotransmitter receptor action (i.e., dopamine, serotonin, N-methyl-D-aspartate) in suicidal behaviour (Reisch et al. 2010), but few have been done to try to understand the functional connectivity of neural circuits in profoundly depressed individuals who express an intention to end their lives. Given the devastating consequences of untreated or treatment-resistant cases of depression, it is imperative that research efforts continue toward developing more

effective treatment and prevention strategies. Up to 15% of patients with treatment-resistant depression die by suicide (Souery, Papakostas, and Trivedi 2006). The aim of the present work is to review recent evidence from neuroimaging studies, lesion studies, connection and network studies, in an investigation of the structural and functional components of the mechanisms involved in suicidal behaviour.

Complexity, Importance, and Uniqueness to Human Life

Some aspects of depression may be represented somewhat analogically through animal models (Malkesman et al. 2009), but even the most sophisticated of these models cannot adequately inform us on the debilitating human psychiatric illness of major depressive disorder (MDD) and its consequences (Anisman and Matheson 2005). At best, these animal models merely give us insight on a small number of depression-relevant behaviours implicated in the illness (Grippo, Cushing, and Carter 2010). Included among these may be learned helplessness (Chourbaji et al. 2005; Seligman 1972), chronic stress (Blanchard, McKittrick, and Blanchard 2001), dysfunctional social behaviours (Pryce et al. 2005), reward prediction (Slattery, Markou, and Cryan 2007), and pain-like behaviours (Blackburn-Munro 2004). However, diverse as they are, animal models are unable to truly provide us with a holistic view of suicidal behaviour.

Despite the fact that these depression-relevant behaviours can be modelled and investigated in animals, the phenomena of suicide and of MDD as an illness, with all associated psychiatric and somatic markers (Penninx et al. 2013), appear to be limited to humans (Prete 2011). In changing significant aspects of behaviour, motivation, cognition, and self-perception (American Psychiatric Association 2000), being affected by MDD unequivocally shrouds all of the characteristics that defines one as human. Based on evidence that has emerged from research in neuroscience, it has become clear that the mechanisms that allow for the perception of self and for engaging in complex and abstract thought may also contribute to what makes the problem of suicide a uniquely human problem (Prete 2011).

Overview of Recent Neuroscientific Research

A wealth of research has focused on investigating risk factors associated with suicidal behaviour (Joiner, Brown, and Wingate 2005; Nock et al. 2008; Pompili 2010). It is evident that there are a number of factors that contribute to the manifestation of suicidal behaviour in patients with MDD. With the aid of recent developments in brain imaging techniques, such as positron emission tomography (PET) and functional

magnetic resonance imaging (fMRI), it has become possible study connectional abnormalities in humans, insofar as correlating activational patterns between brain regions in healthy controls and in MDD patients who have attempted suicide can shed light on the dysfunctional connectivities of these circuits (Anand et al. 2005). In addition, lesion studies (Drevets 2000) can even further strengthen the conclusions drawn from imaging research. This review focuses on two important aspects of human behaviour that become maladaptive in those who engage in or have engaged in suicidal behaviour: (a) the ability to engage in self-referential thought; and (b) the ability to feel complex and abstract emotions. Research suggests that the aberrant connectivity between and within important neural structures plays a major role in the problematically altered perceptions often held by highly depressed individuals, which ultimately contributes a great deal to the expression of the symptoms characteristic of the disorder.

A. Self-Referential Thought and Emotion Regulation

Suicidal behaviour is most commonly implicated in cases in which patients have a history of mood disorders (including both unipolar and bipolar depression), although Shneidman (1993) maintains that the presence of a clinical disorder is not always necessary for suicidal behaviour. Suicidal behaviour is linked to highly negative appraisals of the self (Habenstein, Reisch, and Michel 2013; Meerwijk, Ford, and Weiss 2013), such that the self-concept, in this case pertaining to the process of viewing oneself in a self-referential or phenomenological manner (Northoff et al. 2006), is cognitively distorted. Self-conscious emotions, such as guilt, are considered by emotion researchers (Lewis 2008) to be much more complex, and tend to have greater intensities in individuals with suicidal thoughts or intentions. For instance, these individuals commonly experience feelings of being burdensome on others (Beck and Lester 1976; Ribeiro and Joiner 2009; Van Orden et al. 2006); it would thus be fruitful to examine the neural activity among the cortical networks involved in self-referential thought processes in those individuals who are prone to suicidal behaviour.

Although the regions that are typically activated during processing of emotional stimuli include subcortical structures such as the nucleus accumbens, the amygdala, and the insula (Lemogne et al. 2012), one influential meta-analysis (Northoff et al. 2006) of 27 PET and fMRI studies found that among the areas most prominently activated during self-referential thought are the structures of the anterior cortical midline structures (anterior CMS); namely, the ventromedial prefrontal cortex (VMPFC), the dorsomedial prefrontal cortex (DMPFC), and the pre/subgenual anterior cingulate cortices (pgACC,

sgACC). Structural imaging accounts provide evidence for a reduction in volumetric grey matter particularly in the sgACC across depressive episodes, in both unipolar and bipolar depression (Drevets 2007).

An early model of self-focus proposed by Carver and Scheier (as cited in Lemogne et al., 2012) defined it as the process of adjusting the discrepancy between one's currently perceived state of self and a salient standard to which one holds oneself; depressive moods or symptoms arise when individuals consistently fall short of this and are unable to negate this discrepancy. Given the aberrant functional connectivity between the VMPFC/DMPFC and the amygdala, which receives input pertaining to affective states, patients with MDD are subject to chronically experiencing negative or aversive emotion when engaging in self-referential thought (Lemogne et al. 2012). Evidence suggests that the VMPFC plays an important role in integrating visceromotor aspects of emotional processing by synthesizing environmental and sensory cues, received via the orbitofrontal cortex (OFC), with internal affective states, and in light of this, it is suggested that the VMPFC may even be implicated in the decision-making process (Gusnard et al. 2001). In MDD, abnormalities exist in the networks formed between the MPFC, OFC, and their connected limbic structures, which together form a 'visceromotor network' responsible for modulating emotional behaviour (see Drevets 2007). The medial region of the PFC has been found to have decreased levels of cerebral blood flow (CBF) and glucose metabolism in patients with MDD—and an increase in these levels has been found in patients who have successfully remitted (Klein et al. 2010; Drevets 2007). Neural activity in the limbic system organize outward expressions of behaviour to stressors and emotional stimuli, and so the disconnectivity among components of the anterior CMS with the core limbic structures thus may provide an explanation for why depressed patients, and patients with a history of suicidal behaviour especially, have emotional responses that are incongruent with the nature of stimuli. Based on the conclusions drawn from these imaging studies, it is clear that the integrative networks among the MPFC, the OFC, and the amygdala serve critical functions in both emotional inhibition and expression.

The use of lesion studies is valuable because unlike imaging methods, which provide insight into correlational patterns in brain activity, they can be used to attribute changes in behaviour to certain areas (Koenigs and Grafman 2009). Thus, they can often provide stronger evidence for causal processes involved with certain behavioural expressions. Lesions of the MPFC and OFC have been associated with socially inappropriate behaviour, impulsivity, and lack of concern for moral principles (Ciaramelli et al. 2007). On the other hand, overactivity across these neural networks in patients with MDD could

have implications in the fact that feelings of guilt and burdensomeness is a common characteristic of the disorder.

B. Complex and Abstract Emotions in Perception of Pain

While the amygdala receives affective input, the insula receives interoceptive input (Lemogne et al. 2012), which is implicated in the emotional perception of pain. Results of fMRI studies consistently suggest that two areas that have substantial involvement in pain perception are the anterior cingulate cortex (ACC) and the insula (Zhuo 2006)—areas that, as first described by Broca and then later Papez (as cited in Mayberg et al., 1997), also have critical roles in moderating emotional states. The pathways linking the dorsal and ventral areas of the ACC, which contribute to the normal coordination of autonomic and willed behaviour, motor activity, and mood, are disrupted in patients with MDD (Mayberg et al. 1997). Data from PET scans indicate that the nociceptive pathway also includes the prefrontal cortex (PFC) and the amygdala (Mee et al. 2006); it appears that the only area that is exclusive to pain perception is the somatosensory cortex—all other brain areas have been found to be common to the perception of pain, the regulation of emotion, and the appraisal of self. A surgical procedure known as subcaudate tractotomy disconnects white matter tracts that connect the VMPFC with the subcortical structures has been successful in reducing symptoms of both depression as well as anxiety (Koenigs et al. 2008), further suggesting that it is the excess of neural activity along these connections that contributes to the expression of depressive symptoms.

The term *psychache* is used in the literature to mean a chronic, seemingly endless and inescapable experience of unbearable amounts of intense psychological pain, likely a result of consistent negative and aversive self-appraisals (Mee et al. 2011; Meerwijk and Weiss 2011; Shneidman 1993). Two leading researchers of suicide, Shneidman (1993) and Joiner (2005), have suggested that the most prominent risk factor contributing to suicidal behaviour is the intensity of the psychological, social, or emotional pain experienced by an individual, regardless of whether or not a clinical diagnosis for a depressive disorder exists. Indeed, some of the measure that are used to assess this mental pain are illustrative of this, containing items such as “My pain makes my life seem dark,” “I hurt because I feel empty,” and “My soul aches” (Holden et al. 2001, 225). In a review of psychological pain, Mee and colleagues (2006) compiled a brief list of statements made by patients in description of their personal experiences, for example, “It is like being in a black hole and trying to claw my way up to get out of it but I keep slipping further and further down

that hole. The suffering is torture. It is the worst pain that I know” (Holden et al. 2001, 682).

Incredibly loaded with complex and abstract representations of emotion, these scale items and statements seem to reflect characteristics that are uniquely human. With its connectivity to core structures of the limbic system, the VMPFC plays a key role in making appropriate appraisals and making logical decisions (Ciaramelli et al. 2007) based on information received from emotional processing centres (such as the amygdala), erratic signals sent from these centres is likely to lead to faulty appraisals of the self, of the situation, or of others—congruent with this emotional information.

Lesion studies of the ACC has demonstrated that damage sustained to this area results in a significant increase in response to pain, as well as heightened emotional sensitivity to aversive pain-related memories (Johansen, Fields, and Manning 2001), which may provide an explanation for why depressed individuals experience intense psychological pain. When this pain becomes unbearable, patients are at a much greater risk for committing suicide.

Conclusions and Findings

Given the fact that many of the same structures involved in all of self-perception, perception of the environment, and emotion regulation share communicative networks, it brings together a perspective on the structural and functional mechanisms implicated in MDD and suicide to understand how the connections work together on the scale of the whole brain. As is suggested by Long and colleagues (2015), all of the characteristics that come together to become what is known as a depressive disorder are likely not the result of a dysfunction in any single connection or structural component, but rather, an exceedingly complex interplay of problems that involve several networks and nodes across the entire brain.

Without the use of the neuroimaging and surgical techniques that we have access to, treatments such as deep brain stimulation (DBS) or transcranial magnetic stimulation (TMS) would not be possible. Although methods such as DBS and TMS are usually used only in cases where the depression appears to be resistant to treatment (Schlaepfer et al. 2014), and still bear the chance of relapse, they are inarguably much more effective than last-resort treatment measures taken in past (i.e., induced lesioning in performing lobotomies, or removal of whole brain areas; see Corkin 2002), before advanced neuroimaging technology was available.

Clinical Implications

In the course of a year, suicide accounts for over one million deaths worldwide (Jollant et al. 2010). Understanding the nature of suicide is critical to develop more effective treatment strategies, and one way that this is made possible is through neuroimaging technologies.

The overall benefits of having these advanced neuroscientific techniques is that treatment methods can potentially become more available and accessible to patients who are in need of better paradigms that will assist them in overcoming treatment-resistant depression. Determining a timely, effective line of treatment is critical especially in those patients who are at risk for suicide. However, the problem with more intensive treatment methods is that they can be costly (Hallett 2000), thus making them largely inaccessible. Especially given the consequences of profound, untreated depression, understanding the disorder from both a functional and structural perspective, through the use of neuroimaging, lesion, and network techniques, is a critical step in treating it in the most effective manner possible. However, it is critical to ensure that a balance is achieved in developing these treatments. As is illustrated by studies using populations who had suffered lesions, a complete deactivation of a system of networks can produce the opposite effect, such as aggression a complete deactivation of a system of networks can produce an effect on the polar opposite of the behavioural spectrum, such as the aggression, apathy, callousness, and impulsivity often implicated in fronto-temporal dementia (Ciaramelli et al. 2007). Ultimately, the goal is to strike this balance in allowing patients to be in control of their emotions, but not so much that they hardly feel obligated to adhere to social or moral norms.

Future Directions

It is worth considering that depression as a psychiatric illness may be a complex system of disruptions in functional connectivity, rather than a dysfunction of a single network (Long et al. 2015). It is common for researchers to use a region of interest (ROI) strategy when investigating the neural bases of particular behaviours; however, using this strategy can significantly limit the findings because if researchers decide to examine a particular area a priori, affected structures and networks which are not directly in the locus of the ROI may be missed and a truly holistic view of the system would not be achieved.

One notable question that has been proposed but not yet examined in depth pertains to the differential function of the dorsomedial versus the ventromedial prefrontal cortex.

There have been preliminary studies conducted on these two areas (Koenigs and Grafman 2009; Koenigs et al. 2008) and even in the case that regions are cytoarchitectonically universal, the functions and processes that are carried out by individual neurons may follow a different pattern even in the same conventional brain area. The distinctive functional characteristics between the areas would be worth investigating in greater depth in order to develop solutions to better target these networks.

In closing, it is suggested that, rather than being basic diagnoses that can be identified through confirming items on an inventory of symptoms, major depressive disorder may be viewed as intricate plexa that all interact with one another. Moreover, the characteristics implicated in these complex networks of behaviour are not ones that can be entirely eliminated in the treatment of disorders in which they have gone awry; instead, these are characteristics that must be preserved insofar as they are part of what makes us human. In bringing together convergent evidence using a variety of different techniques, the goal of finally bringing treatments that are more effective to patients is something that current research efforts should aim to achieve.

References

- American Psychiatric Association. 2000. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)*. Arlington, VA: American Psychiatric Association.
- Anand, Amit, Yu Li, Yang Wang, Jingwei Wu, Sujuan Gao, Lubna Bukhari, Vincent P Mathews, Andrew Kalnin, and Mark J Lowe. 2005. "Activity and Connectivity of Brain Mood Regulating Circuit in Depression: A Functional Magnetic Resonance Study." *Biological Psychiatry* 57 (10): 1079–88.
- Anisman, Hymie, and Kim Matheson. 2005. "Stress, Depression, and Anhedonia: Caveats Concerning Animal Models." *Neuroscience & Biobehavioral Reviews, Animal Models of Depression and Antidepressant Activity*, 29 (4–5): 525–46.
- Beck, Aaron T, and David Lester. 1976. "Components of Suicidal Intent in Completed and Attempted Suicides." *The Journal of Psychology* 92 (1st Half): 35–38.
- Blackburn-Munro, Gordon. 2004. "Pain-Like Behaviours in Animals—How Human Are They?" *Trends in Pharmacological Sciences* 25 (6): 299–305.
- Blanchard, Robert J, Christina R McKittrick, and D Caroline Blanchard. 2001. "Animal Models of Social Stress: Effects on Behavior and Brain Neurochemical Systems." *Physiology & Behavior, Social Stress: Acute and Long-term Effects on Physiology & Behavior*, 73 (3): 261–71.
- Chourbaji, S, C Zacher, C Sanchis-Segura, C Dormann, B Vollmayr, and P Gass. 2005. "Learned Helplessness: Validity and Reliability of Depressive-Like States in Mice." *Brain Research Protocols* 16 (1–3): 70–78.
- Ciamelli, Elisa, Michela Muccioli, Elisabetta Lådavas, and Giuseppe di Pellegrino. 2007. "Selective Deficit in Personal Moral Judgment Following Damage to Ventromedial Prefrontal Cortex." *Social Cognitive and Affective Neuroscience* 2 (2): 84–92.
- Corkin, Suzanne. 2002. "What's New with the Amnesic Patient H.M.?" *Nature Reviews Neuroscience* 3 (2): 153–60.
- Dieserud, Gudrun, Espen Roysamb, Oivind Ekeberg, and Pal Kraft. 2001. "Toward an Integrative Model of Suicide Attempt: A Cognitive Psychological Approach." *Suicide and Life-Threatening Behavior* 31 (2): 153–68.
- Drevets, Wayne C. 2000. "Neuroimaging Studies of Mood Disorders." *Biological Psychiatry* 48 (8): 813–29.

- Drevets, Wayne C. 2007. "Orbitofrontal Cortex Function and Structure in Depression." *Annals of the New York Academy of Sciences* 1121 (1): 499–527.
- Grippo, Angela J, Bruce S Cushing, and C Sue Carter. 2010. "Depression-Like Behavior and Stressor-Induced Neuroendocrine Activation in Female Prairie Voles Exposed to Chronic Social Isolation." *Psychosomatic Medicine* 69 (2): 149–57.
- Gusnard, Debra A, Erbil Akbudak, Gordon L Shulman, and Marcus E Raichle. 2001. "Medial Prefrontal Cortex and Self-Referential Mental Activity: Relation to a Default Mode of Brain Function." *Proceedings of the National Academy of Sciences* 98 (7): 4259–64.
- Habenstein, Astrid, Thomas Reisch, and Konrad Michel. 2013. "Brain Regions Associated with Psychological Pain." *Brain Imaging and Behavior* 7 (3): 243–44.
- Hallett, Mark. 2000. "Transcranial Magnetic Stimulation and the Human Brain." *Nature* 406 (6792): 147–50.
- Holden, Ronald R., Karishma Mehta, E. Jane Cunningham, and Lindsay D. McLeod. 2001. "Development and Preliminary Validation of a Scale of Psychache." *Canadian Journal of Behavioural Science* 33 (4): 224–32.
- Johansen, Joshua P, Howard L Fields, and Barton H Manning. 2001. "The Affective Component of Pain in Rodents: Direct Evidence for a Contribution of the Anterior Cingulate Cortex." *Proceedings of the National Academy of Sciences of the United States of America* 98 (14): 8077–82.
- Joiner, Thomas E, Jessica S Brown, and LaRicka R Wingate. 2005. "The Psychology and Neurobiology of Suicidal Behavior." *Annual Review of Psychology* 56 (1): 287–314.
- Jollant, Fabrice, Natalia S Lawrence, Emilie Olie, Owen O'Daly, Alain Malafosse, Philippe Courtet, and Mary L Phillips. 2010. "Decreased Activation of Lateral Orbitofrontal Cortex during Risky Choices under Uncertainty Is Associated with Disadvantageous Decision-Making and Suicidal Behavior." *NeuroImage* 51 (3): 1275–81.
- Klein, Julia, Christine Winter, Nicolas Coquery, Andreas Heinz, R Morgenstern, Andreas Kupsch, and Georg Juckel. 2010. "Lesion of the Medial Prefrontal Cortex and the Subthalamic Nucleus Selectively Affect Depression-Like Behavior in Rats." *Behavioural Brain Research* 213 (1): 73–81.
- Koenigs, Michael, and Jordan Grafman. 2009. "The Functional Neuroanatomy of Depression: Distinct Roles for Ventromedial and Dorsolateral Prefrontal Cortex." *Behavioural Brain Research* 201 (2): 239–43.

- Koenigs, Michael, Edward D Huey, Matthew Calamia, Vanessa Raymond, Daniel Tranel, and Jordan Grafman. 2008. "Distinct Regions of Prefrontal Cortex Mediate Resistance and Vulnerability to Depression." *The Journal of Neuroscience* 28 (47): 12341–48.
- Lam, Ben C P, Michael Harris Bond, Sylvia Xiaohua Chen, and Wesley C H Wu. 2010. "Worldviews and Individual Vulnerability to Suicide: The Role of Social Axioms." *European Journal of Personality* 24 (7): 602–22.
- Lemogne, Cédric, Pauline Delaveau, Maxime Freton, Sophie Guionnet, and Philippe Fossati. 2012. "Medial Prefrontal Cortex and the Self in Major Depression." *Journal of Affective Disorders Journal* 136 (1): e1–11.
- Lewis, M. 2008. "Self-Conscious Emotions: Embarrassment, Pride, Shame, and Guilt." In *Handbook of Emotions*, 742–56.
- Long, Zhiliang, Xujun Duan, Yifeng Wang, Feng Liu, Ling Zeng, Jing-ping Zhao, and Huafu Chen. 2015. "Disrupted Structural Connectivity Network in Treatment-Naive Depression." *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 56: 18–26.
- Malkesman, Oz, Daniel S Pine, Tyson Tragon, Daniel R Austin, Ioline D Henter, Guang Chen, and Hussein K Manji. 2009. "Animal Models of Suicide-Trait-Related Behaviors." *Trends in Pharmacological Sciences* 30 (4): 165–73.
- Mann, J J. 2003. "Neurobiology of Suicidal Behaviour." *Nature Reviews Neuroscience* 4 (10): 819–28.
- Mann, J J, C Waternaux, G L Haas, and K M Malone. 1999. "Toward a Clinical Model of Suicidal Behavior in Psychiatric Patients." *The American Journal of Psychiatry* 156 (2): 181–89.
- Mayberg, Helen S, Stephen K Brannan, Roderick K Mahurin, Paul A Jerabek, Jerold S Brickman, Janet L Tekell, J Arturo Silva, et al. 1997. "Cingulate Function in Depression." *NeuroReport* 8 (4): 1057–61.
- Mee, Steven, Blynn G Bunney, William E Bunney, William Hetrick, Steven G Potkin, and Christopher Reist. 2011. "Assessment of Psychological Pain in Major Depressive Episodes." *Journal of Psychiatric Research* 45 (11): 1504–10.
- Mee, Steven, Blynn G Bunney, Christopher Reist, Steve G Potkin, and William E Bunney. 2006. "Psychological Pain: A Review of Evidence." *Journal of Psychiatric Research* 40 (8): 680–90.

- Meerwijk, Esther L, Judith M Ford, and Sandra J Weiss. 2013. "Brain Regions Associated with Psychological Pain: Implications for a Neural Network and Its Relationship to Physical Pain." *Brain Imaging and Behavior* 7 (1): 1–14.
- Meerwijk, Esther L, and Sandra J Weiss. 2011. "Toward a Unifying Definition of Psychological Pain." *Journal of Loss and Trauma* 16 (5): 402–12.
- Nock, Matthew K, Guilherme Borges, Evelyn J Bromet, Christine B Cha, Ronald C Kessler, and Sing Lee. 2008. "Suicide and Suicidal Behavior." *Epidemiologic Reviews* 30 (1): 133–54.
- Northoff, Georg, Alexander Heinzl, Moritz de Greck, Felix Bermppohl, Henrik Dobrowolny, and Jaak Panksepp. 2006. "Self-Referential Processing in Our Brain: A Meta-Analysis of Imaging Studies on the Self." *NeuroImage* 31 (1): 440–57.
- Penninx, Brenda W J H, Yuri Milaneschi, Femke Lamers, and Nicole Vogelzangs. 2013. "Understanding the Somatic Consequences of Depression: Biological Mechanisms and the Role of Depression Symptom Profile." *BMC Medicine* 11 (1): 129.
- Pompili, Maurizio. 2010. "Exploring the Phenomenology of Suicide." *Suicide and Life-Threatening Behavior* 40 (3): 234–44.
- Preti, Antonio. 2011. "Animal Model and Neurobiology of Suicide." *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 35 (4): 818–30.
- Pryce, Christopher R, Daniela Rüedi-Bettschen, Andrea C Dettling, Anna Weston, Holger Russig, Boris Ferger, and Joram Feldon. 2005. "Long-Term Effects of Early-Life Environmental Manipulations in Rodents and Primates: Potential Animal Models in Depression Research." *Neuroscience & Biobehavioral Reviews, Animal Models of Depression and Antidepressant Activity*, 29 (4–5): 649–74.
- Reisch, Thomas, Erich Seifritz, Fabrizio Esposito, Roland Wiest, Ladislav Valach, and Konrad Michel. 2010. "An fMRI Study on Mental Pain and Suicidal Behavior." *Journal of Affective Disorders* 126 (1–2): 321–25.
- Ribeiro, Jessica D, and Thomas E Joiner. 2009. "The Interpersonal-Psychological Theory of Suicidal Behavior: Current Status and Future Directions." *Journal of Clinical Psychology* 65 (12): 1291–99.
- Schlaepfer, Thomas E, Bettina H Bewernick, Sarah Kayser, Rene Hurlemann, and Volker A Coenen. 2014. "Deep Brain Stimulation of the Human Reward System for Major Depression: Rationale, Outcomes and Outlook." *Neuropsychopharmacology* 39 (6): 1303–14.

- Seligman, Martin E. 1972. "Learned Helplessness." *Annual Review of Medicine* 23 (1): 407–12.
- Shneidman, Edwin S. 1993. "Commentary: Suicide as Psychache." *The Journal of Nervous and Mental Disease* 181 (3): 145–47.
- Slattery, David A, Athina Markou, and John F Cryan. 2007. "Evaluation of Reward Processes in an Animal Model of Depression." *Psychopharmacology* 190 (4): 555–68.
- Souery, Daniel, George I Papakostas, and Madhukar H Trivedi. 2006. "Treatment-Resistant Depression." *The Journal of Clinical Psychiatry* 67 (suppl 6): 16–22.
- Turecki, Gustavo. 2014. "The Molecular Bases of the Suicidal Brain." *Nature Reviews Neuroscience* 15 (12): 802–16.
- Van Orden, Kimberly A, Meredith E Lynam, Daniel Hollar, and Thomas E Joiner. 2006. "Perceived Burdensomeness as an Indicator of Suicidal Symptoms." *Cognitive Therapy and Research* 30 (4): 457–67.
- Zhuo, Min. 2006. "Molecular Mechanisms of Pain in the Anterior Cingulate Cortex." *Journal of Neuroscience Research* 84 (5): 927–33.

Journal of Cognition and Neuroethics

What is Enhancement?

Catherine Gee

University of Waterloo

Biography

Catherine Gee is a PhD student in philosophy at the University of Waterloo in Waterloo, Ontario. Her current work is on the philosophical and ethical issues with enhancement technologies with a special focus on authenticity. Her other primary research interests are ethics, bioethics, and philosophy of psychiatry.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). January, 2016. Volume 3, Issue 4.

Citation

Gee, Catherine. 2016. "What is Enhancement?" *Journal of Cognition and Neuroethics* 3 (4): 41–54.

What is Enhancement?

Catherine Gee

Abstract

Defining enhancement involves the normative concept of 'good health' which is the center of considerable philosophical debate. Determining what counts as a good level of health is required to establish a baseline that anything over and above denotes enhancement territory and anything below would be considered a treatment. I argue the best way to determine health is via a modified version of Wakefield's two criterion for disease. If a case is able to meet both criterion then it can be considered diseased, and thus any intervention involved can be deemed a treatment. However, if a case is unable to meet the two criterion then it is absent disease and therefore healthy and any intervention is thus an enhancement.

Keywords

Enhancement, Treatment, Health, Disease

An enhancement technology is a terminology designated for the things we use to augment or enhance our physical or mental traits. Some are so 'low tech' or common we may not even think of them as enhancements as they are part of our daily lives. Some examples of these more basic enhancements include hair dye, a strong cup of coffee, high heels, and push-up bras. The hair dye allows us to change our hair colour or hide our grays, coffee makes us more alert and awake, high heels make the wearer taller, and push-up bras make breasts appear larger and more lifted. Enhancements can be as simple as the ones listed here or require more complicated technologies such as surgery or psychiatric medication. One of the most frequently used definitions of enhancement is Eric Juengst's which states enhancements are "interventions designed to improve human form or functioning beyond what is necessary to sustain or restore good health" (1998, 29). Defining enhancement involves the normative concept of 'good health' which is the center of considerable philosophical debate. Determining whether a particular intervention qualifies as an enhancement versus a treatment therefore depends, in part, on what constitutes 'good health'. Agreeing on what counts as a good level of health is required to establish a baseline that anything over and above this line denotes enhancement territory. The purpose of this article is to tackle this problem in order to find a useful definition of health (and disease to contrast it with) to establish this baseline upon. In the first section I will discuss how we can determine whether an intervention is better classified as a treatment or an enhancement. The second section

will dig deeper into this distinction by examining three main accounts of health and disease to determine which one provides the best baseline upon which to ground the treatment versus enhancement distinction. Naturalism, normativism, and a hybrid account that blends the two will be discussed and I will argue the hybrid account is the best view for understanding health in a way that allows us to distinguish treatment from enhancement. In the third section I will apply this account of health to some examples to demonstrate how the two criterion in the hybrid account can be used to distinguish treatment from enhancement. This article merely scratches the surface in the conceptual and philosophical issues surrounding enhancement but serves as a starting point to introduce the reader to the technological interventions that are becoming increasingly popular and often surrounded in controversy.

I. What is Enhancement?

Enhancements are often contrasted with treatments and the distinction is grounded in a baseline of good health. Often the technology is the same in both cases, a treatment is intended to restore a patient from below an average level of health to a baseline level of health whereas an enhancement is used to bring a patient to a higher level of health than the baseline. This comparison can be referred to as the treatment verses enhancement distinction and it seems rather intuitive at first glance. For example, consider the use of anabolic steroids in two different individuals, one who has a muscular disorder so the steroids help restore his weakened muscles to a baseline level of health, versus one who is a body builder and uses the steroids to make his already large muscles even bigger and stronger. Another example is the use of Ritalin in an individual with an attention-deficit disorder to restore her attention to a normal level so that she may focus and complete every-day tasks, compared to a student without attention problems who left studying for an exam to the last minute and uses Ritalin as an aid to improve his focus so he can cram for the exam. While the technology is the same in each example – steroids in the first and Ritalin in the second – it is the baseline level of health that determines whether the use of the technology is a treatment or an enhancement (treatments for the individuals with the muscular disorder and attention-deficit disorder, enhancements for the bodybuilder and the student). While these examples make the distinction between treatment and enhancement relatively easy to distinguish not all situations are quite this straightforward, especially in cases of psychological enhancements.

Peter Kramer's account in *Listening to Prozac* of the first patient he treated with Prozac is an example of such a case. Tess was a woman in her early thirties who was

referred to Kramer because she was clinically depressed. After being put on Prozac Tess' life was transformed; she felt rested and hopeful, more relaxed yet energetic, she laughed more frequently, and a new social life bloomed as a result. "She was astonished at the sensation of being free of depression" recalls Kramer (1997, 7). All aspects of her life were changed as a result, and for the better, it seemed. After about nine months Tess went off Prozac and continued to do well. However, after about eight months off medication Tess told Kramer she was slipping and claimed "I am not myself" (1997, 10). Tess was no longer depressed, but wished to take Prozac again to feel as good as she had when she was on it. When Tess first started seeing Kramer she met many of the signs and symptoms of depression such as "tears and sadness, absence of hope, inability to experience pleasure, feelings of worthlessness, [and a] loss of sleep and appetite" (1997, 3). It was clear she was not well and Kramer used Prozac as a means to restore her to health, thus the Prozac was used as a treatment. But what of her request to take it again in the absence of depression, would this qualify as a treatment or an enhancement? Tess asserted Prozac "had lent her surety of judgment; she no longer tortured herself over whether she was being too demanding or too lenient" (Kramer 1997, 9). "It makes me confident" she told Kramer (1997, 9), and it was for these reasons Tess wanted another prescription, not because she was hopeless and not sleeping or eating. It is more difficult to determine whether Tess' request for a second prescription for Prozac qualifies as a treatment or an enhancement than it is in the steroids and Ritalin examples above. This is because a baseline level of health is not an easy concept to define and will vary depending upon which account of health one is using. Thus, before we can establish a baseline we need to determine the best account of health and disease that should be used to build this baseline on. The next section will discuss the three main views of health in the philosophical literature for this purpose.

II. Accounts of Health and Disease

To review, the distinction between treatment and enhancement depends on a clear definition of health, as anything below a standard baseline level of health can be considered a disease in need of treatment and anything above a standard of health would be considered exceptionally healthy and any intervention would be regarded as an enhancement. However, there is little consensus on what exactly "health" entails, as it can be regarded as simply as "freedom from disease" or as broadly as the World Health Organization's notion of "a state of complete physical, mental, and social well-being" (Parens 1998, 3). This section will focus on three main views of health in the

philosophical literature: naturalism, normativism, and a hybrid theory that combines the advantages of both views. Naturalism argues that disease is a bodily malfunction that causes harm, whereas normativism asserts that disease is merely a human construct or value-judgement as there is no biological basis for disease (Murphy 2015, s. 2). Hybrid theories blend the two and assert that the most accurate account of disease is a bodily dysfunction that we disvalue. I will discuss each view and argue that a hybrid view of health is the best view to form a baseline level of health upon which we can distinguish health from disease, and in turn determine whether an intervention is better classified as a treatment or an enhancement.

(a) Naturalism

Naturalists pursue definitions based on the ideal of value-free scientific theory and attempt to highlight that which is biologically natural and normal for a species (Ereshefsky 2009, 221). The focus for this approach is on physiological and psychological states to determine if an organ or bodily system is normal or functioning properly (Ereshefsky 2009, 221). Naturalism is the most prominent view in the literature (Ereshefsky 2009, 222, for examples see Kendell 1975, Scadding 1990, Wachbroit 1994) and within this view Christopher Boorse's work is considered to be the most well-developed and influential. Boorse argues that health is the absence of disease, where "disease" refers to an internal state "that depress a functional ability below species-typical levels" (1969, 542). Thus, for Boorse, "[h]ealth as freedom from disease is then statistical normality of function, i.e., the ability to perform all typical physiological functions with at least typical efficiency" (1969, 542). The benefit of his approach, he argues, is that it removes the problems of normativism entirely as his theory is founded in natural science, not the results of an evaluative decision (1969, 543). The most common objection to Boorse's argument is that as a result of removing values from the equation, naturalism does not properly reflect our use of the terms "health" and "disease" (Ereshefsky 2009, 222). A stock example used to demonstrate this point is homosexuality, as it was only rather recently that the American Psychiatric Association (APA) removed homosexuality from their *Diagnostic and Statistical Manual of Mental Disorders* and thus no longer considered it to be a disease (Ereshefsky 2009, 222). Ereshefsky argues, however, that "[t]he change in classifying homosexuality as a disease was not accompanied by a change in our medical knowledge of homosexuality. What changed, some argue, is whether or not homosexuality is a disvalued state by the APA" (2009, 222). This is the biggest objection

to naturalism, that it is simply impossible to define health or disease without taking values into account¹. As such, this is the launch point for the normativist position.

(b) Normativism

Normativists argue that biological processes alone cannot determine whether or not someone is healthy or diseased, as it is our human values which make this decision (Murphy 2015, s. 2). Unlike the naturalist view which argues if a system is objectively malfunctioning it is a clear indication a disease is present, normativism asserts there are no objective malfunctions. The system or process is judged by us to be unusual or abnormal “because they depart from some shared, usually culturally specific, conception of human nature” (Murphy 2015, s. 2). The physiological or psychological states we desire we label ‘healthy’ and the ones we do not wish to have are called ‘diseased’ (Ereshefsky 2009, 223). Biological systems themselves are neutral, it is the value we attribute to them that determines whether we find them good or bad (Margolis 1976, Goosens 1980, Sedgwick 1982). Engelhardt explains this well in the following passage:

Disease does not reflect a natural standard or norm, because nature does nothing – nature does not care for excellence, nor is it concerned for the fate of individuals qua individuals. Health, insofar as it is to indicate anything more than the usual functions or abilities of the members of the species, must involve judgements as to what members of that species should be able to do – that is, must involve our esteeming a particular type of function. (1976, 266)

Engelhardt is clear that by emphasizing the role value judgements play in our conceptions of health and disease does not deny “that there are real causes of disease or real empirical factors important in maintaining health or causing disease”, just that we need to acknowledge that discussions of health and disease “presupposes evaluations of ourselves and our ambiances” (1976, 267). The problem with Engelhardt’s view is that without natural and objective standards to define disease it can become far too subjective and we are left in the position of not being able to disagree with the claim a particular condition is a disease. Ereshefsky uses an example in the literature about American doctors in the nineteenth century who thought that slaves who attempted to escape suffered from a disease called “drapetomania” (2009, 224). While our contemporary perspective disagrees

1. This objection applies more widely than just normativism – the idea that science itself is value-free is a highly controversial issue.

with this assessment and thinks the attempt to escape slavery and be free can hardly be considered the symptom of a disease, a normativist, claims Ereshefsky, cannot say these nineteenth century physicians were objectively wrong. All Engelhardt and other normativists can argue is that we merely have different values or ideologies than those doctors (Ereshefsky 2009, 224). The problems with this are obvious and subject normativism to issues of relativism by relying on social constructs and values alone. Thus, a better account of health and disease is needed to balance the weight value carries in distinguishing healthy from unhealthy.

(c) Hybrid Theories

There are combinations of naturalist and normativist views called 'hybrid theories' that attempt to use the best features of each view. A main motivation for these theories is to fix the problems with normativism, specifically that normativism allows any undesirable state to be considered a diseased state (Ereshefsky 2009, 224). The hybrid solution is to argue 'disease' is best defined as "disvalued states with the proper biological etiology" (Ereshefsky 2009, 224). A well-known hybrid account is Wakefield's who argues:

A condition is a disorder [disease] if and only if (a) the condition causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criterion), and (b) the condition results from the inability of some internal mechanism to perform its natural function, wherein a natural function is an effect that is part of the evolutionary explanation of the existence and structure of the mechanism (the explanatory criterion). (1992, 384)

Thus the hybrid account has two criterion for disease, a state must be both disvalued and biologically dysfunctional, instead just one criterion as both naturalism and normativism endorse (for another hybrid example see Reznick 1987). As a result hybrid theories narrow the range of cases that qualify as a 'disease' and avoid the counterexamples that render naturalism and normativism problematic (Ereshefsky 2009, 224). However, this account is not exempt from its own problems, namely that due to its restrictive nature it too quickly dismisses controversial cases that we might still want to consider a disease (so we can receive treatment for it) but it may not meet Wakefield's two criterion (Ereshefsky 2009, 224). Ereshefsky uses the example of a woman unable to achieve an orgasm. The function of the clitoris, he explains, is described as providing a female with the capacity to have an orgasm, but this capacity was not selected for via evolution, it

is merely a by-product of selection for the male orgasm² (2009, 224). Thus, a woman's capacity to have an orgasm does not have an evolutionary function which means it fails Wakefield's section criterion, and as such cannot be considered a disease. Ereshefsky argues "[b]ecause Wakefield equates health with no disease, controversial cases fall on the health side of the health-disease dichotomy. A woman's inability to have an orgasm is [therefore, according to Wakefield] a healthy state (no dysfunction)" (2009, 224).

Wakefield argues that the "failure of a naturally selected function is necessary for disorder" and rejects claims that biological disorders can be heterogeneous with evolutionary dysfunctions being only one type of dysfunction (1999, 376). His insistence on the evolutionary basis of the biological dysfunction is to ensure we are looking at the right sort of dysfunctions (ones that are objectively dysfunctional), and not ones that fail to function "in a socially preferred manner" (based on values alone) (Wakefield 1992, 381). By limiting the concept of dysfunction to natural functions selected for via evolution that have gone awry Wakefield attempts to "distinguish dysfunction from other disvalued conditions" (Sadler and Agich 1995, 222). To illustrate this, consider the nose which functions to hold up one's glasses and the heart which functions to pump blood throughout the body (Wakefield 1992, 381). If someone has a nose that is shaped in a way that does not allow it to hold up his eyeglasses it can be said his nose is failing to function in this manner, and if his heart fails to pump his blood properly it too would be failing to function. However, argues Wakefield, we cannot say the oddly-shaped nose that fails to hold up eyeglasses is a nasal disorder, but we would say the heart failing to pump blood is a dysfunctional heart. The difference between the two cases is that holding up eyeglasses is not the *natural* function of the nose (it is merely a way we use it in addition to its actual purpose³, to breathe and smell); whereas the natural function of the heart *is* to pump blood, so when the natural function fails it is said to be disordered (Wakefield 1992, 381-382). In sum, for Wakefield, "[o]nly natural functions are relevant to disorder attributions" (1999, 375).

This argument is far from being widely accepted there are critics who disagree with Wakefield rooting biological dysfunction in evolutionary theory (Bergner 1997, Lilienfeld & Marino 1995, Sadler & Agich 1995). I also disagree with this constraint as there is far too much controversy in determining which biological functions have been naturally selected and which have not that it seems more trouble than it is worth to insist on

2. Note this is a controversial claim.

3. He refers to functions such as these as "intentional uses" (Wakefield 1999, 375).

including evolutionary theory in a concept of health. Furthermore, I think it is overly restrictive and unnecessary for my present purpose. While I agree with Wakefield's motivation to ensure not just any disvalued condition can be considered a disorder, this can be achieved by using his two criterion definition of disorder and removing the evolutionary constraint. To reiterate, Wakefield's two criterion are:

- (a) the condition causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criterion), and
- (b) the condition results from the inability of some internal mechanism to perform its natural function, ~~wherein a natural function is an effect that is part of the evolutionary explanation of the existence and structure of the mechanism~~ (the explanatory criterion). (1992, 384)

I prefer the inclusion of the term "natural function" rather than just using "function" on its own to narrow the types of functions we might consider candidates for biological dysfunction. Back to Wakefield's example of the nose, it seems intuitive that the natural biological function of the nose is to breathe and smell, not to hold up one's eyeglasses. Thus, if there is a dysfunction of a biological mechanism that impedes one's breathing or smelling through their nose (and breathing and smelling through the nose are functions we value) than this dysfunction can be said to be a disease or disorder of the nose, unlike if one could breathe and smell fine but one's nose was shaped in such a way that it failed to hold up one's glasses effectively. Even if we valued noses holding up our glasses, this is not the nose's natural function and as such it is not a disease when it cannot.

In conclusion, despite potential problems with the hybrid account I think the combination of the naturalist and normativist criterion gives us the best shot at a cohesive and practical account of health to base the treatment versus enhancement distinction on. The solution to controversial cases can be resolved using some common sense when using my modified version of Wakefield's two criterion to distinguish disease. These criterion are:

- (a) the condition causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criterion), and

- (b) the condition results from the inability of some internal mechanism to perform its natural function (the explanatory criterion).

If a case meets both of these criterion then it can be said to be properly classified as a disease and if it does not then we can say the individual is healthy. Thus, health can be properly classified as “absence from disease”.

III. Using the Hybrid Account to Distinguish Treatment from Enhancement

Recall from the first section that often the same technological intervention can be used to both treat and enhance depending on the circumstances. How we can distinguish between the two is to establish a baseline level of health that anything beyond that point can be considered an enhancement as it would be making an already healthy individual even healthier or more capable, and anything below can be considered a treatment as it could restore the individual to a baseline level of health. I have argued in section two that health is the absence of disease, and disease can be determined by using a modified version of Wakefield’s hybrid theory of health. Therefore, disease occurs when the condition causes some harm or deprivation of benefit to the person and is the result of the inability of some internal mechanism to perform its natural function. Applying this formula to the example of the use of anabolic steroids in two different individuals can tell us why the use of steroids by the person with a muscular disorder is best classified as a treatment and why the body builder’s use of the same technology is an enhancement. The muscular disorder causes the first individual harm and deprives him of the benefit of properly functioning muscles (criterion one) and his weak muscles are caused by the failure of his bodily mechanisms to perform their natural function (criterion two). As a result of meeting both these criterion the individual can be considered diseased and the anabolic steroids can be used as a treatment to help restore his weakened muscles to a baseline level of health (a state that fails to meet both criterion). The body builder, in contrast, is unable to meet both of the criterion. His large and fully functioning muscles do not cause him harm as valued by society in general, although perhaps the case could be made (particularly within the body building community) that his muscles failing to grow beyond their already above-average size could be considered a harm in virtue of his body-building ambitions. However, the body builder would be unable to meet the second criterion, as his muscles failing to grow as big as he wants them to is not the result of some internal mechanism failing to perform its natural function. His muscles are functioning just fine, significantly above average in fact. As the body builder is unable to

meet both criterion for disease he is therefore healthy, and as such if he used steroids to further increase the size of his muscles it would be an enhancement.

This is a fairly straightforward case, but how effective is the hybrid theory for more complicated ones? Recall the Tess case from the first section. Tess had two main complaints that led to the consideration of prescribing her Prozac by her doctor. Tess was initially referred to him because she met many of the signs and symptoms of depression – she wasn't sleeping or eating well, she felt hopeless, sad, worthless, and was no longer able to experience pleasure. These symptoms meet both criterion for disease. First, it is clear that her symptoms were causing Tess harm for they severely hindered her quality of life from how well she slept and ate to how she felt. Secondly, while there is not enough information in the example for us to know the cause of Tess' depression⁴ there are numerous complicated physical causes of depression that are the result of a biological mechanism failing to perform its natural function. Even if the cause of Tess' depression was not initially physical in nature, the resulting depression does affect her neural chemistry which impedes its natural function (a healthy brain is not a depressed brain). Therefore, I argue that the reason for Tess' first visit to her doctor met both criterion for disease and thus his prescription for and her use of Prozac is best classified as a treatment.

After her depression had been cured and she had been off Prozac for about eight months Tess complained she no longer felt like herself and wanted to feel as good as she had when she was taking the medication. Feeling "good" for Tess was the self confidence she felt when on Prozac, she felt better about herself and the decisions she made. Tess seemed to be unable to feel the level of self-confidence she did when she was medicated. Should a lack of self-confidence be considered a disease? It could be argued that a lack of self-confidence could meet the first criterion for disease, as it is something our society values. We tend to respect those who are self-assured and have more confidence in the decisions made by someone who is sure of herself than in someone who is not. However, a lack of self-confidence is not the result of a biological function gone awry and thus it fails the second criterion. Self-confidence is a psychological trait that exists on a continuum that people possess to varying degrees – some have an over-abundance of self-confidence, some have very little, and most probably fall somewhere in between the two extremes. Simply because our society values self-confidence does not mean that an absence of it is a disease. That is the purpose of the second criterion for disease – to ensure not everything that society disvalues counts as disease. Therefore, Tess' lack of self-

4. Kramer himself may not have known as often clinicians are unable to know the exact cause of mental disorders.

confidence is unable to pass both the value and the explanatory criterion for disease and as such, despite her lack of self-confidence, she is healthy. Since Tess is healthy her second request for Prozac is best classified as an enhancement.

In closing, I have argued in this article that in order to distinguish treatment from enhancement we need to establish a baseline level of health upon which to ground the distinction. Any technological intervention above and beyond this baseline is best classified as an enhancement and any intervention below this baseline should be considered a treatment. The best way to determine health is via a modified version of Wakefield's hybrid theory of disease. In order to be considered a disease the impairment must meet the following two criterion:

- (a) the condition causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criterion), and
- (b) the condition results from the inability of some internal mechanism to perform its natural function (the explanatory criterion).

If an impairment does not meet both criterion then it is best considered to be healthy. Health is defined as the absence of disease and technological interventions aimed at making the diseased healthy are treatments, and interventions aimed at making the healthy even healthier are best classified as enhancements.

References

- Bergner, Raymond. 1997. "What is psychopathology? And So What?" *Clinical Psychology: Science and Practice* 4: 235-248.
- Boorse, Christopher. 1969. "What a Theory of Mental Health Should Be." *Journal for the Theory of Social Behaviour* 6 (1): 61-84.
- Engelhardt, Hugo. 1976. "Ideology and Etiology." *Journal of Medicine and Philosophy* 1 (3): 256-68.
- Ereshefsky, Marc. 2009. "Defining Health and Disease". *Studies in History and Philosophy of Biological and Biomedical Sciences* 40 (3): 221-27.
- Goosens, William. 1980. "Values, Health, and Medicine." *Philosophy of Science* 47: 100-115.
- Juengst, Eric T. 1998. "What Does Enhancement Mean?" In *Enhancing Human Traits: Ethical and Social Implications*, edited by Erik Parens, 29-47. Washington, DC: Georgetown University Press.
- Kendell, Robert. 1975. "The Concept of Disease and Its Implications for Psychiatry". *British Journal of Psychiatry* 127: 305-315.
- Kramer, Peter. 1997. *Listening to Prozac*. New York: Penguin Group Inc.
- Lilienfeld, Scott, and Lori Marino. 1995. "Mental Disorder as a Roschian Concept: A Critique of Wakefield's "Harmful Dysfunction" Analysis." *Journal of Abnormal Psychology* 104: 411-420.
- Margolis, Joseph. 1976. "The Concept of Disease". *The Journal of Medicine and Philosophy* 1: 238-255.
- Murphy, Dominic. 2015. "Concepts of Disease and Health." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. <http://plato.stanford.edu/archives/spr2015/entries/health-disease/>.
- Parens, Erik. 1998. "Is Better Always Good? The Enhancement Project." In *Enhancing Human Traits: Ethical and Social Implications*, edited by Erik Parens, 1-28. Washington, DC: Georgetown University Press.
- Reznek, Lawrie. 1987. *The Nature of Disease*. London: Routledge & Kegan Paul.
- Sadler, John, & George Agich. 1995. "Diseases, Functions, Values, and Psychiatric Classification." *Philosophy, Psychiatry, and Psychology* (2): 219- 231.
- Scadding, John. 1990. "The Semantic Problem of Psychiatry." *Psychological Medicine* 20: 243-248.

Sedgwick, Peter. 1982. *Psychopolitics*. New York: Harper and Row.

Wakefield, Jerome C. 1992. "The Concept of Mental Disorder. On the Boundary Between Biological Facts and Social Values." *The American Psychologist* 47 (3): 373–88.

Wakefield, Jerome C. 1999. "Evolutionary Versus Prototype Analysis of the Concept of Disorder." *Journal of Abnormal Psychology* 108 (3): 374-399.

Wachbroit, Robert. 1994. "Normality as a Biological Concept." *Philosophy of Science* (61): 579–591.

Journal of Cognition and Neuroethics

Does the Human Right to Health Include a Right to Biomedical Enhancement?

Martin Gunderson
Macalester College

Biography

Martin Gunderson is DeWitt Wallace Professor of Philosophy at Macalester College where he teaches courses on human rights and normative ethics. His primary research interests are social philosophy and bioethics.

Acknowledgements

An earlier version of this paper was presented May 22, 2015 at the Neurotherapeutics and Psychosurgery Conference sponsored by the Center for Cognition and Neuroethics in Flint, Michigan.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). January, 2016. Volume 3, Issue 4.

Citation

Gunderson, Martin. 2016. "Does the Human Right to Health Include a Right to Biomedical Enhancement?" *Journal of Cognition and Neuroethics* 3 (4): 55–75.

Does the Human Right to Health Include a Right to Biomedical Enhancement?

Martin Gunderson

Abstract

If we grant that there is a human right to health then we are committed to a human right to biomedical enhancement. In particular, I argue that the human right to health should be interpreted to include biomedical enhancements within its scope in the sense that there is a limited liberty right to pursue biomedical enhancements and a rights-based justification for limited entitlements to biomedical enhancements. I begin with a discussion of the human right to health in international law and practice and assume for the sake of argument that the legal human right to health is morally justified. After discussing the human right to health in international law, I argue that the underlying functions that we value when we value health are scalar and do not provide a threshold between therapy and enhancement. I go on to consider various principles philosophers and policy analysts have used to apply the human right to health equitably. None of principles provides a threshold between therapy and enhancement. I end by suggesting that if there is a moral human right to health it too must include biomedical enhancements within its scope.

Keywords

Biomedical Enhancement, Health, Human Rights, Right to Health

Introduction

Much of the work on human rights and biomedical enhancement has argued that various aspects of biomedical enhancement pose dangers that require the protection of human rights or even that human rights themselves are threatened by biomedical enhancements. George Annas et al, for example, argue that there should be an international treaty prohibiting germ-line genetic engineering (Annas, Andrews and Isasi 2002, 151-178). Article 13 of the European Council's Convention on Human Rights and Biomedicine states, "An intervention seeking to modify the human genome may only be undertaken for preventive, diagnostic or therapeutic purposes and only if its aim is not to introduce any modification in the genome of any descendants" (European Council 1999). Notwithstanding such skepticism about biomedical enhancement, I argue that biomedical enhancement falls within the scope of the human right to health. If we grant the human right to health, then we are committed to a human right to biomedical enhancement. In particular, I argue that the human right to health should be interpreted

to include biomedical enhancements within its scope in the sense that there is a limited liberty right to pursue biomedical enhancements and a rights-based justification for limited entitlements to biomedical enhancements.

The right to biomedical enhancement is a limited right. It is certainly true that biomedical enhancement could be used in ways that violate human rights, but this is true of a variety of human rights. Important activities protected by human rights such as speech, religion, and participation in government can be used to violate rights. The human rights that protect such activities need to be limited and balanced with other rights. This is no less true of the human right to biomedical enhancement.

I begin with a discussion of the human right to health in international law and practice and assume for the sake of argument that the legal right to health is morally justified. It is worthwhile beginning with international law because of the degree to which a right to health has been worked out and put into practice by international organizations such as the World Health Organization (WHO). After discussing the right to health in international law, I argue that the underlying functions that we value when we value health are scalar and do not provide a threshold between therapy and enhancement. I go on to consider various principles philosophers and policy analysts have used to apply the human right to health equitably. None of principles provides a threshold between therapy and enhancement. Although I do not assume that the legal human right to health mirrors a moral human right to health or even that there is a moral human right to health, I end by suggesting that if there is a moral human right to health it too must include biomedical enhancements within its scope. In the end, we are better off considering biomedical interventions on a case-by-case basis without worrying about whether the intervention is therapy or enhancement.

The Right to Health in International Law

Article 12.1 of the International Covenant on Economic, Social, and Cultural Rights (ICESCR) provides for "...the right of everyone to the enjoyment of the highest attainable standard of physical and mental health" (UN General Assembly 1966b). The human right to health is a complex right with several component rights that present different challenges for financing health care and limiting the scope of the general right to health. These component rights include a liberty right to pursue health, a socioeconomic right to guaranteed access to health-related goods and services, and a right not to be subject to discrimination on grounds of race, ethnicity, gender, religion, or national origin in the

distribution of health-related goods and services. The component rights can be partially understood in terms of the corresponding duties they impose on states.

For the interpretation of Article 12 it is helpful to turn to General Comment 14 of the UN Committee on Economic Social and Cultural Rights as well as various UN resolutions and declarations.¹ General Comment 14, which has been particularly influential, provides that states have duties to respect, protect and fulfill the right to health (CESCR 2000, Paragraphs 33-37). States have a duty to respect the right to health in part by not interfering with attempts by individuals to provide for their health. In addition, states have a duty to protect individuals from coercive interference with the enjoyment of the right to health by third parties. These duties create a liberty right to pursue one's health. In this regard, General Comment 14 states, "The right to health contains both freedoms and entitlements. The freedoms include the right to control one's health and body..." (CESCR 2000, Paragraph 8).

The duty to fulfill the right to health requires states to ensure access to health services, and this threatens to create enormous budgetary pressure. Fortunately General Comment 14 takes account of this and distinguishes two sorts of duty regarding the fulfillment of the right to health. States have a core obligation to ensure primary health care, including the provision of adequate food to prevent hunger, adequate shelter, essential drugs as defined by WHO, immunization against common childhood diseases, and safe water regardless of budgetary constraints (CESCR 2000, Paragraphs 43-44). States cannot justify non-compliance with these core obligations on financial grounds (CESCR 2000, Paragraph 47). Beyond this basic duty states under budgetary pressure are to progressively realize the fulfillment of the right to health. It creates, in effect, a goal that states are obligated to pursue within reasonable budgetary constraints.

Article 12 of the ICESCR does not spell out what is meant by health, and we need to turn to documents such as the Alma Ata Declaration, the World Health Organization Constitution, and General Comment No. 14. These documents provide two characterizations of health. The broadest and most controversial is the WHO definition, which defines health in terms of complete physical, psychological, and social wellbeing (WHO 1946). The WHO definition gained influence when it was codified in the Alma

1. While the general comments of UN treaty committees, resolutions and declarations are not binding international law they do carry legal weight because they are often cited by lawyers in international tribunals and influence adjudication and state practice as well as the practice of UN agencies such as the WHO. They can also evolve into international customary law as they are adopted by state practice and acknowledged at least implicitly by states as legally authoritative. As a result, these instruments are often referred to as soft law, as opposed to binding international law (Blake 2008).

Ata Declaration, which was adopted shortly after the ICESCR went into effect (WHO 1978, Article 1). A second approach is exemplified by General Comment 14, which notes that Article 12 of the ICESCR did not adopt the WHO definition (CESCR 2000, Paragraph 4). Although General Comment 14 does not give a specific definition of health, its explanation of what counts as a violation of the duty of states to respect the right to health makes it clear that health should be characterized in terms of preventing bodily harm and unnecessary morbidity and mortality. Paragraph 50 of General Comment 14 states, “Violations of the obligation to respect [the human right to health] are those State actions, policies or laws that contravene the standards set out in article 12 of the Covenant and are likely to result in bodily harm, unnecessary morbidity and preventable mortality” (CESCR 2000).

The Value of Health and Biomedical Enhancements

It is unfortunate that the meaning of “health” is not spelled out more clearly because whether the human right to health provides for entitlements to biomedical enhancements and a liberty right to pursue enhancements depends on how health is characterized. If one follows ordinary usage and regards health as the absence of disease, disability and psychological disorder (i.e., the absence of pathology), then it is obvious that the scope of the human right to health does not include biomedical enhancements.² Enhancements by definition go beyond what is necessary to cure or prevent disease and disorder. This is also true of Norman Daniels’ characterization of health as species typical functioning (Daniels [1985] 2008, 37). Enhancements aim at improvement over species typical functioning or normality.

If, on the other hand, the WHO definition of health is adopted, there is a straightforward argument for including biomedical enhancements within the scope of the right to health. Since, on the WHO definition, the human right to health protects complete physical, mental and social wellbeing, it clearly includes enhancements. This is true whether one adopts an objective list account of wellbeing or a subjective account. Subjective accounts of wellbeing characterize wellbeing in terms of mental states such as satisfied preferences or pleasurable states. If the right to health is taken to include satisfied preferences, for example, it is clear that biomedical enhancements are within its scope. Objective list theories characterize wellbeing in terms of states that make a person’s life better even if the person does not desire or prefer them. Relevant examples

2. Norman Daniels is right that it is closer to ordinary usage to characterize health as the absence of pathology rather than merely the absence of disease (Daniels [1985] 2008, 36).

might include clarity of mind and having temperaments that help one to act virtuously. If health includes such things, then biomedical enhancements are also included along with therapies insofar as they can increase mental or physical qualities that contribute to objective wellbeing.

The WHO definition is open to plausible counter examples, however. If wellbeing is taken to mean what philosophers often mean—the extent to which a person’s life is going well for that person—then it is not a necessary condition of health. A person could lack complete wellbeing because the person is unhappy or has an inadequate standard of living and yet be healthy (Bognar and Hirose 2014, 31). Unless wellbeing is characterized in terms of an objective list that includes health, it is also arguable that complete wellbeing is not sufficient for health. A person could have complete wellbeing in terms of preference satisfaction or pleasurable states and yet be disabled or unhealthy because of an undetected disease (Houseman 2006, 254).

Even if we reject the WHO definition of health, however, it still makes sense to say that in valuing health we value more than the mere absence of pathology. As various writers have noted, pathology undermines valuable human physical and psychological functioning (Daniels [1985] 2008, 37; Yamin and Norheim 2014, 30). I shall call these “health-related functions” and speak in terms of health-related functioning. These are the functions for which biomedical interventions can be relevant. But, in valuing health-related functioning we value more than mere species typical functioning (normality) or the absence of pathology. We value being as high functioning as possible. Depression, for example, tends to undercut one’s motivation and rob one of vitality. In valuing vitality, however, we do not simply value normal vitality. Increases in vitality above the normal level are also valuable. Buchanan et al get at this when they note that Prozac was originally used solely as an anti-depressant, but was eventually also prescribed to make non-depressed people feel better. They claim that what people care about is whether a biomedical service is beneficial and affordable, not whether it cures disease (Buchanan, Brock, Daniels and Wikler 2000, 98). The same is true for physical conditions. Loss of vision is disvalued in large part because it impairs a variety of functions that we can accomplish with vision such as easy mobility within our physical surrounding. Once again, however, enhanced vision would add value. The things we value in valuing health have positive scalar values that do not establish a threshold at the absence of disease and infirmity or at species typical functioning. Of course, there is a maximal level of functioning beyond which further visual acuity, for example, is counter-productive, but that level is not at the level of the absence of disease and disability or species typical functioning. Health-related functions are valuable whatever account we give of the

ultimate importance of health. Whether, for example, health is necessary for fair equality of opportunity, as Norman Daniels claims, or for wellbeing health-related functions are a valuable component (Daniels [1985] 2008, 42-46). Call this the “expanded notion of health.” On the expanded notion of health, health includes the degree of function one has regarding health-related functioning (those functions that are threatened by pathologies such as disease, disability and psychological disorder).

The expanded notion of health fits well with the capabilities approach developed by Amartya Sen and Martha Nussbaum (Nussbaum 2000, 2011; Sen 1985, 1992). A person has health-related capabilities by having access to what is needed to attain a physical and psychological functioning of the sort disease and infirmity undermine. As Sen and Nussbaum note, the capabilities approach provides for individual freedom to choose whether to pursue various functions (Sen 2004, 334). The right to health does not require states to provide people with health-related functions. Rather what the right to health requires is *access* to what is necessary for those functions. Put in terms of the capabilities approach, it covers capabilities for health in the expanded sense. Nussbaum claims that human rights generally provide for basic capabilities and that insofar as human rights are respected by states they can be analyzed in terms of capabilities (Nussbaum 2002, Sec. 4; Sen 2005). According to Nussbaum, appealing to human rights is a way of making justified claims to treatment respecting one’s basic capabilities (Nussbaum 2002, 138-139). On Nussbaum’s approach, it could be argued that human rights protect capabilities and that the human right to health protects capabilities relevant to health. Characterized in this manner, the human right to health includes both biomedical therapies and biomedical enhancements within its scope. Both are means to attaining high levels of physical and psychological functioning.

The expanded notion of health also fits with ordinary language. We often use the word “health” to describe the state of being free of pathology, and this produces counter-examples to the WHO definition of health. Note, however, that it also makes sense to say that someone is extremely healthy or super healthy. Such a person has a high degree of health-related functions.

Emphasizing the expanded notion of health is respectful of individual autonomy. Individual autonomy is best characterized as control over one’s health and body limited by risks posed to others. Individuals exercising the right to liberty component of the human right to health might well decide to opt for safe and effective means of biomedical enhancement should they become available in the future. Note also that the value of health for a person depends in part on the person’s other values and life projects (Broome 2002, 95). Complete health for an athlete requires access to different medical

treatments than full health for a monk. This is part of the justification for the stress international health agencies place on participation in the adoption and application of the human right to health. The expanded notion of health takes account of individual autonomy by giving people a greater range of access to biomedical interventions in order to control their health and bodies.

In addition, the expanded notion of health does not rely on a shaky conceptual distinction between therapy and enhancement. It is notoriously difficult to draw a clear conceptual distinction between biomedical enhancements and therapies. Just when does the use of antidepressants or growth hormones shade off from therapy into enhancement? There are, of course, paradigm cases of enhancement and therapy, and in a rough and ready way we can continue to speak of therapies and enhancements. Setting a broken tibia is clearly therapy while blood doping to increase cycling performance is clearly enhancement. Yet, the distinction lacks the clarity to be a basis for policy in the distribution of health-related services or the adoption of prohibitions on biomedical enhancements generally. If an expanded notion of health is accepted, the good news is that we do not need to worry about precise definitions of biomedical enhancement and biomedical therapy because the distinction is not normatively relevant.

How does the expanded notion of health relate to the human right health? Human rights protect valuable interests, and the human right to health protects the value of health-related functions. Since the functions that give health its value are scalar, biomedical interventions that improve those functions are within the scope of the right to health whether or not they go beyond the prevention of disease and disorder and hence count as enhancements. Ultimately the right to health protects what enables health-related functions. Given the values protected by the right to health it would be arbitrary to limit the scope of the right to health to prevention or cure of pathologies or to species typical functioning. What is plausible about the WHO definition of health is not that complete wellbeing is a necessary and sufficient condition of health, but that health should not be restricted merely to "...the prevention of disease and infirmity" (WHO 1946). In short, the expanded notion of health does not draw a distinction between therapy and enhancement.

Limits to the Human Right to Health

There are, of course, limits. With several exceptions such as the human right to be free from torture and the human right against slavery, states may derogate or limit human rights under certain conditions. Article 4 of the International Covenant on

Civil and Political Rights provides that with certain restrictions human rights in that covenant may be derogated in national emergencies that threaten the nation as a whole, and international customary law has extended this to include public health emergencies (WHO 2005; UN Economic and Social Council, 1985, sec. IB, iv). International customary law also provides that states may derogate rights when essential to maintain respect for the fundamental values of the community (UN Economic and Social Council, 1985, sec. IB, v). None of these limits provides a reason for drawing a line between biomedical enhancement and therapy. States may prohibit putative therapies as well as enhancements that are clearly ineffective or that are dangerous without overriding benefits. In addition, prohibitions essential to maintain respect for fundamental community values might rule out some therapies such as xenotransplants from the great apes as well as some enhancements such as the blood doping of athletes.

Moreover, regulations to prohibit enhancements in general, as opposed to specific enhancements, are likely to be over-inclusive in that they would prohibit medical interventions that are justified. Enhancements ranging from plastic surgery for cosmetic purposes to dental braces are biomedical enhancements, though we would not be justified in prohibiting such practices.³ This sort of over-inclusiveness results in a violation of the right to health, since it does not provide adequate reason for derogating the right to health in such cases.

The most severe constraints on the implementation of the human right to health are budgetary. Norman Daniels makes a good point when he states in *Just Health: Meeting Health Needs Fairly* that we cannot infer specific healthcare entitlements from a human right to health (Daniels [1985] 2008, 15 and 317). Other than the minimum core of primary health services specified by General Comment 14, this is certainly true. Even in the case of life-saving therapies such as a pancreas transplant, to use Daniels' example, it does not follow from the human right to health that one is entitled to a pancreas transplant (Daniels [1985] 2008, 317). Healthcare entitlements depend, at least in part, on the ability of states to finance them. Beyond the nonderogable core obligations, states need to adopt principles of distributive and procedural justice to prioritize various health interventions to determine which ones will be adopted as entitlements.

Since prioritizing is required because of budgetary constraints, an obvious approach is to use a cost-effectiveness analysis to determine what health-related entitlements are necessary in order to satisfy the duty of progressive realization of the human right to

3. The use of braces for cosmetic purposes is Daniel Tobey's example, though he defends distinguishing therapy and enhancement for the purpose of regulating genetic enhancements (Tobey 2003-2004, 158).

health. A common way to do this is to measure health outcomes in terms of quality-adjusted life years (QUALYs) that take account of both the quality of life before and after a medical intervention and the number of additional years of life that can be secured by the intervention. First, the analysis determines quality of life from 0 (death) to 1 (full health) that an intervention will likely secure. The quality of life at issue can be determined through surveys and public discussions. One way of doing this is to use the standard gamble approach and ask people what risk of death they would accept if it meant a possible cure of their disease or disability. Another way is the time trade-off approach that asks how many years of life a person would sacrifice for a treatment that cured a disease or disability. The QUALYs associated with a particular pathology are then compared with the QUALYs of the health state after a therapeutic intervention to determine how many QUALYs a therapy will provide. The final result of the QUALY analysis is the product of the number assigned to the quality of life secured by a particular type of intervention and the number of years that intervention will add to life. Medical interventions can then be ranked on the basis of cost per QUALY.

QUALY analysis is typically used when the issue is cost-effective treatments of pathologies, but QUALY analysis can be applied more broadly to include biomedical enhancements. What constitutes maximal health-related functioning can be characterized in terms of the level of functioning that could be achieved by adopting safe and effective biomedical interventions including enhancements. In the case of biomedical enhancements a state of health without disease or disability could be compared with an enhanced state to determine the number of QUALYs produced by the enhancement.

Although I am focusing on QUALY analysis as a cost-effectiveness tool for the application of human rights, a similar expansion could be used in the case of cost-effective analysis in terms of disability-adjusted life years (DALYs). DALY analysis provides a way of assigning a numerical value to the number of years lived at a certain level of disease or disability. The cost per DALY averted can then be determined. Although DALY analysis is currently used to measure the burden of disease or disability for a person or society, DALY analysis can be expanded to take account of the expanded notion of health. DALY analysis requires setting a base point for life expectancy and a way of determining disease or disability burden. Usually the life expectancy of the nation with the highest life expectancy is used. If the expanded notion of health is adopted, however, life expectancy could be set in terms of life expectancy that would result from the use of safe and effective biomedical interventions including enhancements. The degree of burden can be determined by comparing the present state to an ideal of full health, which can be characterized as maximal functioning possible with safe and effective biomedical

enhancements. This can be done with the same methods used for QALY analysis. Note that, on the expanded analysis, “disease” and “disability” are not the best terms to refer to states that are simply sub-optimal because they are not enhanced.

Since the right to health requires progressive realization of maximal health-related functioning, neither QALY nor DALY analysis will draw a sharp line between biomedical enhancements and therapies. Some biomedical enhancements may be more likely to increase health-related functioning on a cost-effective basis than some therapies and may have priority some biomedical enhancements over some therapies. This might result in such biomedical enhancements becoming entitlements. This is because some enhancements may be more likely to increase health-related functioning at a cost-effective basis than some therapies. Safe, effective and moderate memory enhancement may eventually be more cost-effective and produce more QALYs than aromatherapy, for example. What matters is the equitable distribution of biomedical interventions ranked in terms of their effect on the quality of life relative to the number of years of life added by the intervention. Each sort of intervention needs to be evaluated on its own terms. Although I have focused on QALY and DALY analysis because of their common use, a similar argument could be given for any sort of cost-effectiveness analysis.

It should be noted that QALY analysis has been subjected to a variety of objections. These include claims that QALY analysis discriminates against persons with disability and the elderly, is overly subjective and even arbitrary, and confuses preferences with values (Harris 1987; Daniels and Sabin 2002, Chapter 3). It is not my purpose, however, to defend QALY analysis, but to show that its adoption, as an example of cost-effectiveness analysis, does not justify excluding biomedical enhancements from the scope of the human right to health.

Alternatives to Cost-Effectiveness Analysis

Deliberative Democratic Process: It might be argued that deliberative democratic processes constrained by principles of distributive justice should be the primary method of applying human rights to health when there are budgetary constraints, and Daniels adopts a version of this approach that he calls “accountability for reasonableness” (Daniels [1985] 2008, Chapter 10). The central element of accountability for reasonableness is a process of fair deliberation that requires policies to be adopted on the basis of rationales that are publicly accessible and reasonable in the sense that they appeal to “evidence, reasons and principles that are accepted as relevant by (‘fair minded’) people who are disposed to finding mutually justifiable terms of cooperation” (Daniels 2008, 118).

In addition, the policies adopted must be open to revision, and the process must be governed by public regulation (Daniels [1985] 2008, 118; Daniels and Sabin 1997, 322-343). Note, however, that once the expanded notion of health is adopted along with a right to fulfillment of health, reliance on accountability for reasonableness will not distinguish enhancement from therapy. It is easy to imagine, for example, a fair deliberative procedure resulting in the outcome that life-extending enhancements should be adopted as an entitlement. Unless enhancements are ruled out prior to using the fair deliberative process, as Daniels does, it cannot be assumed that such a process will distinguish enhancements and therapies (Daniels [1985] 2008, 149-155).

Dignity-Based Sufficiencyarianism: In light of the financial concerns generated by the expanded notion of health and its incorporation into the human right to health, it might be objected that we would be better off adopting a view that the human right to health requires only the minimum of health care necessary for a life worthy of human dignity (Nickel [1987] 2007, Chapter 9). Since on this interpretation the human right to health guarantees only what is sufficient for a life worthy of human dignity and no more, I will follow philosophical usage and refer to this as the sufficientarian interpretation of the human right to health.

Sufficientarianism has several advantages. It provides a way of limiting healthcare expenses by the state at a time of tightening budgets. States need only guarantee access to the minimum level of health care specified by the right. In addition, the human right to health gains strength because it can be rigorously enforced without making it virtually impossible for the state to pursue other goals. Basing sufficientarianism on what is necessary for a life worthy of dignity also seems plausible because human rights covenants that specify the right to life are based on human dignity, though the meaning of dignity is not spelled out (UN General Assembly 1948, Preamble; UN General Assembly 1966a, Preamble and 1966b, Preamble).

It might also be claimed that the sufficientarian alternative presents a problem for the thesis I have defended since biomedical enhancements are not necessary for a life worthy of human dignity. On the surface, pain and suffering seem to undermine a life worthy of dignity in a way that forgoing biomedical enhancements does not. Moreover, the moral and legal distinction between therapy and enhancement is intuitively appealing on the ground that alleviating the suffering and incapacity caused by disease and disability should be given priority over the joys of enhanced health-related functioning.

The problem with this objection is that there are therapies that ought to be within the scope of the human right to health even if they are not necessary for a life worthy of human dignity, and there can be enhancements that ought to be covered because they

are necessary for a life worthy of human dignity. Being a person with a disability, for example, is certainly compatible with human dignity, but access to therapy to alleviate a disability is clearly within the scope of the right to health. This is also true of various conditions such as mild to moderate arthritis that are painful but nonetheless compatible with a life worthy of human dignity.

Access to some enhancements might also become necessary to ensure dignity. A life worthy of dignity requires the capabilities necessary to participate in society as a free and equal person. This includes being able to compete for positions in society as well as take part in political processes and being free from oppressive discrimination and prejudice. As Dan Wikler has argued, radical increases in intelligence through enhancement by the majority of a population may adversely affect the requirements for civic participation and hence the equal status of those who are not enhanced; and, as Christine Overall has argued, enhancement might subject already marginalized groups to increased discrimination and prejudice (Overall 2009, 327-340; Wickler 2009, 352). Hence the adoption of biomedical enhancements by some, especially those who are already privileged, in the exercise of the liberty right to health creates a strong reason for guaranteeing access to such enhancements for all citizens when it is necessary to ensure equality under law. If, however, equal access cannot be guaranteed, this constitutes grounds for restricting the liberty right regarding those enhancements. In general, entitlements to enhancements result from the way in which component rights of the right to health interact with one another and with other human rights.

Opportunity-Based Sufficiency: It might be thought that the problems noted above could be avoided by spelling out dignity in terms of having access to a certain range of opportunities or capabilities. This is the approach taken by Norman Daniels who argues that justice requires that people be treated equally in the sense of having access to the normal opportunity range presented by their society (Daniels 2008, Chapter 2). The normal opportunity range is the range of opportunities afforded by a society to persons on the basis of their ability (Daniels [1985] 2008, 43-44). Health care, according to Daniels, is special in terms of justice because its goal is to restore people with pathologies to typical species functioning (normality), and this is necessary for having access to the normal range of opportunities (Daniels [1985] 2008, 44-46).

Attempting to restrict this version of sufficiency to health defined in terms of pathology also conflicts with the liberty right to health combined with the right to equal respect regarding health policy. The liberty right to health, as noted, protects the interest people have in controlling their health and bodies. It follows that people have a right within limits to use biomedical enhancements to enhance their health and bodies

even when it is not necessary to prevent disease or disorder. Moreover, if some people use their liberty to pursue enhancements that give them more opportunities treating people equally may require that all have access to the enhancements. In short, if the liberty aspect of the right to health allows people to engage in biomedical enhancements, this will affect what counts as fair equality of opportunity.

Prioritarianism: It is plausible to suppose that with limited healthcare budgets we should give priority to those who have the most serious health problems. Prioritarianism in the allocation of healthcare resources assigns weight to recipients of healthcare resources in proportion to the severity of their health conditions. If prioritarianism is accepted for the allocation of health care it might be thought that enhancements would be excluded from the scope of a right to health care because they would carry little weight in the allocation process. Although prioritarianism gives more weight to those with severe health problems, it does not follow that no weight is given to health-related concerns addressed by biomedical enhancement. Moreover, some enhancements might be so significant that they would outweigh some therapies even in prioritarian terms. An enhancement that would significantly extend longevity might receive greater weight than therapies for minor ailments. In fact, those who did not receive the enhancement for longevity could be regarded as worse off than those who did. As people exercise their health-related right to liberty to enhance themselves, the unenhanced become worse off.

Buchanan's Enhancement Enterprise

The analysis I have offered provides a reason for accepting what Allen Buchanan calls the "enhancement enterprise." The enhancement enterprise, according to Buchanan, allows considerable freedom to develop enhancement technologies and devotes significant public resources to research on enhancement technologies and policies for coping with enhancements (Buchanan 2011, Chapter 2, especially 60-63). Buchanan defends the enhancement project on pragmatic grounds based on the benefits of enhancement for both individuals and society, while I provide a specific reason for the enhancement project based on the human right to health.

Buchanan also restricts the enhancement enterprise to liberal democratic societies because of his concern that other societies will abuse enhancements (Buchanan 2011, 63). This seems to follow from Buchanan's pragmatic defense of the enhancement enterprise. A defense of enhancements in terms of human rights, however, does not limit enhancements to liberal democracies. Instead the analysis I have offered would limit the enhancement project in a different way. All states are under a duty to pursue the

health of their citizens in accord with the human right to health, and this duty includes biomedical enhancements as well as therapies. All of these interventions, however, need to be pursued within a human rights framework that includes rights that protect individuals from various forms of oppression and inequality. The human right to health is best conceived as part of a suite of human rights, all of which are necessary for wellbeing. In short, all states are under a duty based on the human right to health to pursue the enhancement enterprise in the context of the relevant rights necessary for the protection of individuals.

Is the Human Right to Health a Genuine Right?

Those who see rights as trumps along the lines in which Ronald Dworkin has analyzed U.S. Constitutional rights might object that the human right to health is not a genuine right at all (Dworkin 1977, xi). On Dworkin's analysis, rights trump considerations of welfare. If this is applied to the human right to health it raises an obvious problem since increasing wellbeing and hence welfare is a central component of the right to health. When rights are analyzed as trumps this component of the human right to health looks more like a mere policy goal than a right. Even the liberty component of the right to health is problematic. In human rights law the liberty right to control one's health does not trump considerations of welfare, but can be derogated by considerations of public health and even fundamental community values.

Although the right to health is not a right in the sense in which rights trump all considerations of welfare, it is nonetheless a right and not merely a policy goal. If not trumps, human socioeconomic rights such as the right to health are what James Nickel calls high priority norms (Nickel [1987] 2007, 41). The human right to health requires states to prioritize the wellbeing of their citizens over other objectives not grounded in human rights with a few exceptions for limitations that are generally consistent with wellbeing and hence with the right to health. Also, human rights, unlike mere policy goals, empower citizens to demand state action to provide for the interests protected by human rights. The right to health is no exception. While the right to health does not entail a particular treatment, it nonetheless empowers citizens to demand that the state create institutions that provide for health, protect the health of citizens, and create a reasonable minimum level of health care within the constraints of available resources (Gunderson 2011, 49-62).

The liberty to control one's health care presents more of a challenge since the right can be derogated by fundamental community values. John Harris, for instance, argues for

a human right to reproductive liberty that includes the use of enhancement technologies and bases this on a version of Mill's harm principle that he calls "the democratic presumption" (Harris 2007, 72-79). According to Harris, "only serious real and present danger either to other citizens or to society is sufficient to rebut this presumption. If anything less than this high standard is accepted, liberty is dead" (Harris 2007, 27). On Harris's view, permitting derogation of the liberty right to health on the basis of community values would completely eviscerate the right.

The community values limitation is best interpreted by claiming that the community values that could justify a restriction of liberty are themselves constrained. Article 4 of the International Covenant of Civil and Political Rights, which provides for the derogation of the human right to health, states that no derogation may "...involve discrimination solely on the ground of race, colour, sex, language, religion or social origin." The Siracusa Principles also state that the margin of discretion left to the states to limit derogable rights does not apply to the rule of non-discrimination (UN Economic and Social Council, 1985, Part IB, sec. 28). Hence restricting the liberty right to health to prevent enhancement on the ground that it violates religious percepts of the community, for instance, would not be justified. Harris has a good point, however, concerning the extent to which deference to community values and majority rule can threaten the right to liberty. In light of this concern the relevant community values must be fundamental in the sense that they are a component of the identity of the community and therefore vital for the preservation of the community. In short, the community values exception should be regarded as agreeing with Harris's democratic presumption with the caveat that one of the ways in which society can be threatened by serious real and present danger is to have the values on which society depends undermined. This is no different from various other rights. Germany, for instance, respects the freedom of expression while prohibiting the advocacy of Nazism. Another example would be the use of medical enhancements to create people with abilities so superior and a temperament so aggressive that they threaten to undermine democracy and respect for individual rights.

A Moral Right to Health

I have assumed that the human right to health embodied in international law and practice is morally justified and argued that it is best interpreted as containing a right to biomedical enhancement within its scope. The same argument could be used to show that if it is assumed that there is a moral human right to health it should also be interpreted as containing a right to biomedical enhancement. The scope of a moral

human right to health will be determined in large part by the nature of health, and, as I have argued, there is good reason to adopt what I have called the expanded notion of health. The expanded notion of health, whether legal or moral, contains health-related enhancements including biomedical enhancements. An analysis of the moral right to health will also need many of the features of the legal human right to health. It will, for example, need to be a socioeconomic right that functions as a high priority norm rather than a right that trumps reasons based on human welfare. The human right to health is, after all, a right that seeks human welfare. The moral human right to health, like the legal human right to health, will also compete with other moral human rights and therefore be limited in various ways.

In addition, the moral human right to health, like the legal right to health, will be subject to budgetary constraints. Defenders of a moral human right to health might rely on various strategies to limit the moral human right to health such as cost-effectiveness, sufficientarianism, prioritarianism, democratic deliberation or even a combination of some of these. None of these analyses, as I have argued, provides a reason to create a threshold between biomedical enhancement and biomedical therapy. As a result, the moral human right to health will not draw a sharp line between biomedical enhancements and therapies any more than does the legal human right to health.

Conclusion

At first glance it sounds outlandish to maintain that the human right to health entails a right to biomedical enhancement. It would seem to justify the demand that the state provide access to whatever transformative enhancements a person wants regardless of cost. Once the nature of the human right to health and the constraints on its application are taken into account, however, the claim is far less troubling than it first appears. Moreover, safe, effective, and affordable biomedical enhancements that are transformative are still somewhere in the future.

There are also good reasons to think that the claim that the human right to health includes biomedical enhancements is justified. One reason is based on the value of health. Philosophers have offered a variety of justifications of a right to health ranging from utilitarian considerations of the importance of health for human wellbeing to contractarian considerations of the importance of health for fair equality of opportunity. All of these require placing value on health-related functions, and consequently there is good reason to characterize health in terms of health-related functions. Whatever its ultimate philosophical justification a human right to health protects these functions.

Since these health-related functions are scalar and can be improved by biomedical enhancements as well as biomedical therapies, a human right to health is best understood as including biomedical enhancements within its scope.

Another reason is based on the principles of justice that might be used to apply the human right to health under budgetary constraint. Biomedical enhancements are included within the scope of the human right to health whether we adopt cost-effectiveness principles such as QALY analysis, or principles based on deliberative democracy, sufficientarianism, or prioritarianism. In the end, the human right to health does not make or justify a normative distinction between enhancement and therapy.

References

- Annas George J., L. B. Andrews, and R. M. Isasi. 2002. "Protecting the Endangered Human: Toward an International Treaty Prohibiting Cloning and Inheritable Alterations." *American Journal of Law and Medicine* 28 (2-3): 151-178.
- Blake, Conway. 2008. "Normative Instruments in International Human Rights Law: Locating the General Comment." Center for Human Rights and Global Justice Working Paper Number 17, available at <http://chrgj.org/documents/normative-instruments-in-international-human-rights-law-locating-the-general-comment/> [accessed June 12, 2015].
- Bognar, Greg and Iwao Hirose. 2014. *Health Care Rationing: an Introduction* London & New York: Routledge.
- Broome, John. 2002. "Measuring the Burden of Disease by Aggregating Well-Being." In *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*, edited by Christopher J.L. Murray et al, 91-113. Geneva: World Health Organization.
- Buchanan, Allen. 2011. *Beyond Humanity?* Oxford and New York: Oxford University Press.
- Buchanan, Allen, Dan W. Brock, Norman Daniels and Daniel Wikler. 2000. *From Chance to Choice: Genetics & Justice*. Cambridge: Cambridge University Press.
- Committee on Economic, Social and Cultural Rights (CESCR), General Comment No. 14, (2000), Paragraphs 33-37, E/CN.4/2000/4.
- Daniels, Norman. (1985) 2008. *Just Health: Meeting Health Needs Fairly*. 2nd Edition. Cambridge: Cambridge University Press.
- Daniels, Norman and James Sabin. 1997. "Limits to Health Care: Fair Procedures, Democratic Deliberation, and the Legitimacy Problem for Insurers." *Philosophy & Public Affairs* 26 (4): 305-350
- Daniels, Norman and James Sabin. 2002. *Setting Limits Fairly: Can We Learn to Share Medical Resources?* New York: Oxford University Press.
- Dworkin, Ronald. 1977. *Taking Rights Seriously*. Cambridge MA: Harvard University Press.
- European Council. 1999. Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine:

Convention on Human Rights and Biomedicine. Article 13 (ETS No. 164), entered into force December 1, 1999.

Gunderson, Martin. 2011. "Does the Human Right to Health Lack Content?" *Social Philosophy Today* 27: 49-62.

Harris, John. 2007 *Enhancing Evolution: The Ethical Case for Making Better People*. Princeton & Oxford: Princeton University Press.

Harris, John. 1987. "QUALYfying the Value of Life." *Journal of Medical Ethics* 13 (3): 117-123

Hausman, Daniel. 2006. "Valuing Health." *Philosophy & Public Affairs* 34 (3): 246-274.

Nickel, James. (1987) 2007. *Making Sense of Human Rights*. 2nd Edition. Malden, MA and Oxford: Blackwell Publishing.

Nussbaum, Martha. 2002 "Capabilities and Human Rights." In *Global Justice and Transnational Politics*, edited by Pablo De Greiff and Ciaran Cronin, 117-149. Cambridge MA: MIT Press: 117-149

Nussbaum, Martha. 2011. *Creating Capabilities: The Human Development Approach*. Cambridge, MA and London: Harvard University Press.

Nussbaum, Martha. 2000. *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge U. Press, 2000.

Overall, Christine. "Life Enhancement Technologies: the Significance of Social Category Membership," In *Human Enhancement*, edited by Julian Savulescu and Nick Bostrom, 327-340. Oxford: Oxford University Press.

Sen, Amartya K. 2005. "Human Rights and Capabilities." *Journal of Human Development* 6 (2): 151-166.

Sen, Amartya K. 1992. *Inequality Reexamined*. Oxford: Clarendon Press.

Sen, Amartya K. 1985. "Well-being, Agency and Freedom: The Dewey Lectures 1984." *Journal of Philosophy* 82 (4): 169-221.

Tobey, Daniel L. 2004. "What's Really Wrong with Genetic Enhancement: A Second Look at Our Posthuman Future." *Yale Journal of Law and Technology* 6 (2003-2004): 54-158

United Nations, Economic and Social Council, *The Siracusa Principles on the Limitation and Derogation Provisions in the International Covenant on Civil and Political Rights*. UN Document E/CN.4/1985/4 Annex (1985), Sec. iv.

- UN General Assembly, International Covenant on Civil and Political Rights (ICCPR). 1966a. United Nations, Treaty Series, vol. 999, p. 171.
- UN General Assembly, International Covenant on Civil and Political Rights (ICCPR). 1966b. International Covenant on Economic, Social and Cultural Rights (ICESCR). 1966b. United Nations, Treaty Series, vol. 993, p. 3, Article 12.
- UN General Assembly, International Covenant on Civil and Political Rights (ICCPR). 1948. Universal Declaration of Rights. Universal Declaration of Human Rights, 10 December 1948, 217 A (III).
- Dan Wikler, Dan. 2009. "Paternalism in the Age of Cognitive Enhancement: Do Civil Liberties Presuppose Roughly Equal Mental Ability?" In *Human Enhancement*, edited by Julian Savulescu and Nick Bostrom, 341-355. Oxford: Oxford University Press.
- World Health Organization (WHO). 1978. Declaration of Alma Ata. International conference on primary health care, Alma-Ata, USSR, 6-12 September 1978 Geneva: WHO, 1978.
- World Health Organization (WHO). 1946. "Preamble." The Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.
- World Health Organization (WHO). 2005. *Revisions of the International Health Regulations, Endorsed by the Fifty-eighth World Health Assembly in Resolution 58.3*. Geneva, Switzerland: World Health Assembly 3/05/2005.
- Yamin, Alicia Ely and Ole Frithjof Norheim. 2014. "Taking Equality Seriously: Applying Human Rights Frameworks to Priority Setting in Health." *Human Rights Quarterly* 36: 296-324.

Journal of Cognition and Neuroethics

Resource Egalitarianism, Prosthetics, and Enhancements

Rhonda Martens

University of Manitoba

Biography

Rhonda Martens is an Associate Professor at the University of Manitoba. Most of her work has been on the history and philosophy of astronomy, including the book, *Kepler's Philosophy and the New Astronomy* (2000). She has recently developed an interest in the ethics of cyborg technologies and is co-authoring a science fiction novel with her husband, Rodrigo Muñoz.

Acknowledgements

This paper was originally presented at the Neurotherapeutics and Psychosurgery conference in Flint, Michigan. Thanks to Jami Anderson, Neil McArthur, Meena Krishnamurthy, Simon Cushing, Catherine Gee, Sruthi Rothenfluch, Martin Gunderson, and the other attendees.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). January, 2016. Volume 3, Issue 4.

Citation

Martens, Rhonda. 2016. "Resource Egalitarianism, Prosthetics, and Enhancements." *Journal of Cognition and Neuroethics* 3 (4): 77–96.

Resource Egalitarianism, Prosthetics, and Enhancements

Rhonda Martens

Abstract

Theories of economic justice are concerned with how to divide up resources in a way that is fair. Many resource egalitarians believe that we should divide up resources unequally to compensate for certain types of differences in abilities between individuals. For example, resources may be required to retrofit old buildings with visual fire alarm systems for people who are deaf. Improvements in prosthetics and enhancement technologies (e.g., drugs to increase alertness or memory) could be used to address inequalities in abilities directly. The abilities we have are, to some extent, a matter of chance. Only some people have the ability to become musical virtuosos. Only some people have the ability to become Olympic athletes. Developments in prosthetics and enhancement techniques (e.g., drugs to increase alertness) promise to change chance to choice. As our ability to control our abilities increases, we need to consider whether this is a good idea. We also need to ask what our society should look like in order to promote a beneficial use of new technologies. These issues are complicated by the fact that many of our theories of justice rely heavily on the distinction between chance and choice. We hold people responsible for things they can control, and not for the things they cannot. This article considers a popular theory of distributive justice – Dworkin’s resource egalitarianism – and shows how it gives us the wrong answers to these questions. The problems raised by considering Dworkin’s theory points in the direction of what an adequate theory of justice needs to look like in order to accommodate developments in prosthetic and enhancement technologies.

Keywords

Human enhancement, Prosthetics, Justice, Dworkin

Introduction

Egalitarian theories of economic justice often concern themselves with the problem that individuals come into the world with different sets of abilities (natural resources). Some sets of natural resources are more advantageous in the social (including economic) world than others. Rawls referred to this as the natural lottery. Egalitarians like Dworkin and others propose that social resources be unequally distributed to offset inequalities in natural resources. The job, then, is to figure out how this distribution is supposed to work, which is a complicated and controversial affair. Buchanan, Brock, Daniels, and Wikler have proposed that prosthetic and human enhancement technologies have the potential to allow us to equalize natural inequalities directly, bypassing some of the controversies

involved in social resource distribution (2001, 69). It might seem that this solution would be a welcome one to resource egalitarians like Dworkin. But here's what Dworkin has to say about the possibility of being able to equalize natural inequalities directly:

That crucial boundary between chance and choice is the spine of our ethics and our morality, and any serious shift in that boundary is seriously dislocating. (2000, 444)

We are entitled to think that our most settled convictions will, in large numbers, be undermined, that we will be in a kind of moral free-fall, that we will have to think again against a new background and with uncertain results. (2000, 446)

Dworkin later clarified that the "moral free-fall" is a problem for how we apply our theories of justice to specific cases, but is not a problem for theories of justice generally (2004, 363). For example, if we hold people morally responsible only for outcomes they can control, and if scientific advances changes what we can control, then we will be responsible for more. The rule, 'hold people responsible for what they can control' has not changed, even though the cases to which it applies has.

I will argue that Dworkin is mistaken, that revisions to his theory are necessary not just because of the possibility of future technologies, but on the basis of currently available ones as well. Furthermore, it is my position that examples drawn from prosthetic and enhancement technologies are interestingly different from the usual examples used to motivate resource distribution principles. Exploring these examples will be instructive for determining the shape that we want our theory of resource justice to look like. I'll begin by discussing the role that the distinction between chance and choice plays in Dworkin's theory, as well as the intuitions that motivate this distinction. Next I will turn to an example of a currently existing technology, the cochlear implant, and how this connects with Dworkin's view. I'll briefly discuss a few other examples as well. Considering these cases brings up the issues of health care rationing, accessibility, and responsibility. I'll close with a discussion of what a theory needs to look like that takes these issues into account.

Option and Brute Luck

Dworkin's distribution principle relies on a distinction between option and brute luck, as follows:

Option luck is a matter of how deliberate and calculated gambles turn out . . . Brute luck is a matter of how risks fall out that are not in that sense deliberate gambles . . . Obviously the difference between these two forms of luck can be represented as a matter of degree, and we may be uncertain how to describe a particular piece of bad luck. (1981, 293)

For example, being born to parents who gamble away the grocery money is a matter of bad brute luck for the child, but bad option luck for the parents. For another example, a farmer may choose to plant a certain crop, and then have the bad luck of the weather disagreeing with that choice. This is a combination of option and brute luck. Brute luck, good or bad, can have social as well as natural causes. Dworkin lists being the victim of racism as an example of socially caused brute bad luck (2000, 445-446). For Dworkin, people are to be compensated for the results of bad brute luck but not bad option luck. We can see that the distinction between chance and choice plays a crucial role in his redistribution principle, and so moving the line between chance and choice will change how redistributions pan out.

Part of what motivates the distinction between brute and option luck is an interest in having a theory of distributive justice track our notions of responsibility. Some of the examples Dworkin uses – reckless gamblers, spendthrifts, lazy people – express a concern with free riders, people who do not contribute sufficiently to society but undeservedly benefit from it. Notice the language Dworkin uses in the following: “Why should the spendthrift be rewarded for hard work and frugality he never practiced, out of taxes raised from those who have in fact worked hard and been frugal?” (2002, 113).

A similar motivation might account for Dworkin’s answer to the “equality of what?” question. Dworkin proposes that we equalize resources. Others propose that we equalize wellbeing or welfare (e.g., Griffin 1986). Dworkin believes that welfare egalitarianism requires transfers that “would strike most people as unjust” (2004, 340). Consider the case of the lazy person who wishes also to be wealthy. This is an expensive preference. If we satisfy this preference, we do so at the expense of those who are not lazy or who do not wish to be wealthy. Dworkin’s preferred response to the problem is as follows: “Most lazy people have not chosen to be lazy, but they are free to overcome their laziness, even though they must sometimes make extra effort at the cost of “welfare” to do so” (2003, 193). As long as the lazy person was not deprived of resources, the fact that the resources are insufficient to meet his expensive preferences is not unjust according to Dworkin.

I'm going to make some assumptions for the purposes of this paper. I'm going to assume that we are all, if only for the sake of argument, sympathetic to the idea of resource egalitarianism. Namely, we ought to use social resources to compensate for inequalities in natural resources. I'm also going to assume that we are all, for the sake of argument, sympathetic to the idea that we are not obliged to carry the free rider. Finally, I'm going to assume a garden-variety notion of free will, where sense can be made of the idea that we can make choices at all and can be held responsible for those choices. The goal is to see the extent to which someone holding these assumptions can handle prosthetic and enhancement technologies. That said, it is also worth asking whether these are assumptions we should be holding, even for the sake of argument, but that is a subject for another paper.

One more preliminary matter before we get our examples. Anderson (1999, 295) raised a now well-known objection called the Harshness Objection. The basic idea is this. If someone freely chooses to drive irresponsibly, and freely chooses to not get medical insurance, then according to Dworkin's theory, we are not obliged to provide her with medical treatment if she gets into an accident. This seems harsh. There are several possible responses to the harshness objection. One is to simply bite the bullet by agreeing that it is harsh but that we are still not obliged to provide medical treatment. I'll call this the 'cruel but fair' response (and it is possible that nobody in the literature holds this position). Another response is to propose that we can give people moment-to-moment fresh starts (Fleurbaey 2005, 2008). The reckless driver made a bad decision in the past. We can forgive that, and work our distribution principles from this new moment, resetting the clock. A third is to point out that while we are not obliged to provide medical treatment, we can choose to do so anyways, for some other reason that has nothing to do with our distribution principle (e.g., charity or a trumping principle).

Prosthetics and Enhancements

In this section we'll discuss prosthetics and enhancements in very general terms, but later we will discuss the cochlear implant in a bit more detail. The details matter for justice considerations. For our present purposes, what prosthetics and enhancements have in common is that they can change an ability set that an individual has. I follow Silvers (1998, 101) in distinguishing between modes, levels, and functions of abilities.¹ For example, someone who is paraplegic can travel (function) using a wheelchair (mode)

1. Thanks to Catherine Gee for bringing Silvers' work to my attention.

at a certain pace (level). In the interests of space I will ignore the controversies around specifying normal functions, modes, and levels.

Our social structures and socially created physical structures are designed with certain assumptions about functions, modes, and levels. For example, it is assumed that we all have a flying disability, and so stairs are put in buildings with more than one story. There is a cost to retrofitting buildings and social structures to accommodate functions, modes, or levels that were not taken into account during the original construction. For example, one measure that could be taken towards integrating the Deaf and hearing communities would be to teach nearly everyone sign language. This would be a substantially expensive undertaking at this point in time (teacher retraining, a linguistic gap between those who grew up under the new educational system and those who did not, etc.), but would not be expensive to maintain once fully in place. Often the (initially) cheaper option involves training and using medical devices on individuals to normalize function, mode, and level as much as possible (see Silvers 1998 for a discussion of the dangers of normalization). The costs in this option are often individual rather than societal costs.

The resource egalitarian holds that these individual costs need to be compensated because it is a matter of brute luck that the individual needs to carry these costs in the first place. The problem I will discuss is what happens a prosthetic is introduced that produces (at least near) species-typical functions, modes, and levels. Let's consider a "magic" prosthetic, a future version of the cochlear implant that allows people who are deaf to hear at the same level as someone with species-typical hearing. Once that prosthetic is in place (successful adjustment after surgery), there are no further extra costs to the individual. He is no longer at a disadvantage in the social and economic marketplace. At that point, he no longer needs the compensatory resources reserved for people who are deaf.

So far resource egalitarian gives us intuitive answers for the person who is deaf as a matter of brute luck, and the person who exercises option luck to no longer be deaf. But there's at least one more possibility: the person who exercises option luck to remain deaf by refusing the cochlear implant. On Dworkin's resource egalitarianism, we are only entitled to compensatory resources in the event that our *brute* luck is bad. We are not entitled in the event that our *option* luck is bad. Prosthetics and enhancements change brute luck into option luck (at least to the extent that the prosthetic or enhancement works really well), so on Dworkin's view, the introduction of a new prosthetic or enhancement is an introduction of a new reduction in compensatory resources for those who opt out. The cochlear-refuser may still need the compensatory resources that she was entitled to before the invention of the implant, but now she is no longer entitled

to them. This seems harsh.² Furthermore, the cochlear-refuser is worse off than the person who is not a viable candidate for the cochlear implant. Deafness is still a matter of brute luck for the non-candidate, and so he is still entitled to compensation. Yet it is still the case that both of them are at a disadvantage in a society that assumes a hearing population. Dworkin's theory of distributive justice increases the difficulty of an already difficult decision. It reduces welfare, and in the case of prosthetics, it reduces welfare for an already vulnerable population.³

Perhaps Dworkin can insist on continuing to supply resources on the grounds of a cost-benefit analysis. There are always risks and costs associated with surgery and implants. Surely the risks and costs themselves can be compensated. However, surgical risks are only risks for those who choose to get the surgery. So this provides no grounds for supplying resources to the person who chooses to not get the surgery.

One thing that could be said is that being faced with this choice at all is a case of brute bad luck, and the choice is something for which we should compensate. After all, a hearing person doesn't need to consider whether or not to get the cochlear implant. This seems a possible middle ground, but it isn't clear to me that Dworkin's theory gives us the right results. Presumably it is less bad luck to be faced with the choice than to have no choice at all (the extra bad luck of losing compensation entitlements is introduced by Dworkin's theory!). So the invention of the cochlear implant would still reduce the amount of entitled compensation.

Let's consider a different kind of case. Pre-natal screening can sometimes allow parents to choose termination or continuation of a pregnancy in the event that the foetus is likely to have a costly condition. Suppose the parents choose to continue the pregnancy. The brute luck belongs to the child (and so the child is entitled to compensation if the brute luck is bad), but the option luck belongs to the parents. No compensation is owed to them, even though it will cost more to raise the child. This will put economic pressure on the parents to terminate.

2. Interestingly, the Canadian Academy of Audiology lists "a desire to be a part of the hearing world" as one of the criteria for being a viable candidate for the cochlear implant. <https://canadianaudiology.ca/consumer/cochlear-implants.html> [last accessed April 27, 2015].

3. It is worth pointing out that if a caretaker is refusing a cochlear implant on behalf of a child, then the nature of the luck for the child is brute. That child would still be entitled to compensatory resources. The problem arises when adults make decisions on behalf of themselves.

Suppose pre-natal interventions are available to produce genetic advantages beyond our current species-typical functions, modes, or levels (enhancements). If this becomes a widespread practice, parents who opt out of these interventions could find themselves with extra child-rearing costs (e.g., different schools, increased length of economic dependence due to a marketplace that assumes an enhanced workforce). Again, Dworkin's account does not give us the resources to justify compensating these parents because their luck was option. This puts economic pressure on the parents to enhance.

Physical Autonomy and Coercion

It seems pretty clear that two of the standard responses to the harshness objection misdescribe the cochlear example. It might be cruel but fair to let an unlucky gambler lose her home, but it seems just cruel to remove resources for a cochlear-refuser. The 'fresh start' view also runs into problems. The loss of a house through gambling is a single event in time. It makes sense to consider the possibility of a fresh start from that moment. But the choice of getting the cochlear implant remains an ongoing choice for someone who does not yet have it. Fleurbaey's (2005, 2008) fresh start view does not give unlimited chances. In the reckless gambler case, the gambler gets a fresh start only if she is committed to not gambling anymore.

One move Dworkin could make would be to allow his brute/option luck distinction to be trumped in certain circumstances. The cases we are considering (prosthetics and enhancements) involve modifying the body. Physical autonomy is very important, and is a viable candidate for a trumping consideration. We might be willing to agree that it is usually problematic to coerce someone to get a medically unnecessary surgery. Withholding resources on the grounds that a surgery would make those resources unnecessary seems coercive.

The coercion solution needs further work to help us out (more work than we have space for here). There is a sense in which coercion is unproblematically used on a regular basis (e.g., raising children or threatening potential criminals with incarceration). I'll follow Wertheimer (1987) and Ryan (1980) in viewing coercion as problematic if it involves violating a right or entitlement. On this view, if the cochlear-refusing person is being denied resources to which she is *not* entitled (which, on Dworkin's view, she wouldn't be), then we haven't established that coercion is a problem. But I'm quite uncomfortable with this result.

It's instructive to look at other cases of medical coercion. Recently California SB 277 was passed that requires school children to be vaccinated unless they have a medical exemption.⁴ Personal beliefs against vaccination practices do not count as a medical exemption, and so these personal beliefs can only be exercised if the child is home schooled. It is true that parents who do not wish to vaccinate their children are under tremendous social pressure. It is also true that it is a better thing to have the vaccines available than to relieve the social pressures on those who wish to opt out.

Interestingly, most of the rhetoric justifying the law focuses on the health consequences of the increasing numbers of unvaccinated children rather than on the economic costs of an outbreak.⁵ The claim to an obligation to vaccinate is based on potential harm done to others (e.g., the loss of herd immunity). This consideration does not apply to the cochlear implant. Deafness is not contagious. The issue instead has to do with our economic obligations, and it is harder to make the case that economic considerations trump physical autonomy.

Harshness and Health Care Rationing

I've suggested that it is harsh to deny compensation to the cochlear-refuser or the parent who refuses to enhance his child. But is it unjustly harsh? Perhaps the cochlear-refusers are analogous to those who are reckless with their health, thereby costing society money through their carelessness.⁶ "One man's freedom in health is another man's shackle in taxes and insurance premiums" (Knowles, 1977, quoted in Wikler 2004, 111). If we think in terms of the intuitions that motivate Dworkin's resource egalitarianism, the health-reckless are not entitled to the same coverage (although they might receive it for other reasons).

What makes harshness excessive? Some societies are harsh by necessity. Indeed, all societies are harsh to some extent. The rationale behind health care rationing is that our resources are always finite, and the money used to fund one medical project is money not used to fund another. This is an unfortunate fact, and we want to make these hard

4. http://www.leginfo.ca.gov/pub/15-16/bill/sen/sb_0251-0300/sb_277_bill_20150219_introduced.html [last accessed August 10, 2015.]

5. The CDC website states that "Vaccination Protects Your Family, Friends, and Community." <http://www.cdc.gov/features/ReasonsToVaccinate/> [last accessed August 10, 2015.]

6. Thanks to Sruthi Rothenfluch for raising this question.

decisions in the most ethical manner possible. Dworkin's distinction between brute and option luck gives us one way of providing a principle for how health care gets rationed.

Several critics of resource egalitarianism (e.g., Anderson 1999, Denier 2005) have pointed out practical problems as well as problems with what we want our relationship with the medical profession and the state to be like. It is intrusive and judgemental for a doctor to grill us on our lifestyles before deciding on whether to provide subsidized treatment (or any treatment at all). It is impractical (and intrusive and judgemental) for bus drivers to determine whether a blind person is not at fault for being blind before allowing the guide dog on the bus. Health care rationing is both an ethical and a practical problem, so it needs to be feasible. Rather than pursuing these points further, I will focus on disanalogies between those who refuse prosthetics/enhancements and those who are reckless with their health. We'll focus on the cochlear implant to make these differences clear.

One difference that is particularly relevant to how Dworkin discusses option luck is that the cochlear-refuser knows with certainty that he will remain deaf. He is hoping for a different compensation protocol. The reckless person, by contrast, is aiming big (and often failing). In the case of the gambler, the hope is a large sum of money for a small amount of effort, thereby bypassing the usual, more labour-intensive methods of making money. In the case of the heavy smoker, the hope is a lifestyle that defies probability. Nicotine is pleasurable. But it is a rare smoker who hopes to become sick (and there are more rapid and reliable ways to achieve this goal). Dworkin holds that we should be responsible for our gambles in order for our choice to gamble to be respected (1981, 294). Respecting and supporting liberty is an important part of Dworkin's theory. The issue for the cochlear-refuser is that the choice to refuse now comes with a socially imposed penalty, one that did not exist prior to the existence of the implant. Later we will discuss the significance of the social nature of the penalty.

Another difference has to do with our feelings about the characters used in examples. Anderson (1999), Denier (2005), Wikler (2004) and others point out that when justifying restrictions on health coverage to the reckless, examples are used that pander to our judgmental attitudes. Our society has harsh judgements on addicts, gamblers, and reckless drivers.⁷ The feeling is that they are behaving carelessly and maybe even callously

7. Ubel et. al. (1999) did a study that suggests that some people would deny health care to addicts regardless of whether the addiction caused the health problem or influenced the prognosis. Instead, it seems that some would deny health care on the grounds that addicts are not the sort of people worth saving. A troubling result indeed.

with our hard-earned resources. The example would not be so compelling if we instead pointed out that the risk of breast cancer increases significantly with the choice to delay having children, as well as the choice not to have children at all. It seems fairly clear from this consideration that we view delaying having children as a socially respectable choice that does not involve the careless treatment of group resources (Wikler (2004) uses this example). Furthermore, we would not want delaying or avoiding having children to be a choice available only to the financially well off. Similarly, refusing the cochlear implant or an enhancement also seems a respectable choice that does not involve carelessness. But on Dworkin's view, women who delay having children, cochlear-refusers, and gamblers are all making choices that increase the odds of having bad luck, and thus get treated the same.

Earlier I proposed that we assume for the sake of argument that we are not obliged to carry free riders. I'll follow Arneson (1982, 621-622) in defining free riding as involving a certain kind of reasoning, either explicitly or implicitly. The free rider observes that she will benefit more from a cooperative scheme if she does not contribute, and this observation is her reason for not contributing. While it is certainly possible that a cochlear-refuser might engage in this form of reasoning, it is not a necessary form of reasoning to come to the conclusion to refuse the cochlear. There are plenty of other reasons the refuser might have, and these reasons seem substantially more compelling than the free rider form of reasoning. Similarly, it is possible that a woman might decide to delay children *because* she'll be better off economically if she is childless *and* has health insurance to cover the increased breast cancer risk. It's possible, but would be a very surprising reason for making that decision. So the brute/option luck distinction allows us to avoid carrying the free rider, but it cuts a lot of other people off from compensation as well. It is too strict.

At least some of the considerations pertaining to health care rationing are beside the point. Perhaps the most important difference between the cochlear-refuser and the health-reckless is that while a smoker with lung cancer is sick, being deaf is not an illness. The resources in question are not for treating an illness, but rather are for navigating a society that assumes we can all hear. The issue is accessibility and discrimination rather than health care rationing. Granted, some of the same considerations for health care rationing hold for accessibility. Societies have limited resources, so complete accessibility for everyone may be economically unfeasible.

Accessibility and Direction of Responsibility

There is a growing philosophical literature on accessibility (e.g., Kelly 2013, Toboso 2011, Crossley 2003). I just want to focus on one accessibility issue here by making the following claim: to the extent to which social injustices rather than physical misfortunes render an ability set disabling, there is a stronger demand on society to take responsibility. It seems likely that Dworkin would agree with this claim. Consider what he says about other examples of socially caused brute bad luck:

We feel a greater responsibility to compensate victims of industrial accidents and of racial prejudice, as in both cases victims, though in different ways, of society generally, than we feel to compensate those born with genetic defects or those injured by lightning or in those other ways that lawyers and insurance companies call “acts of God.” (Dworkin 2000, 445-446)

Compensation is backwards-looking (compensation is for an event that occurred in the past), but it seems a friendly amendment to suggest that we also have a greater responsibility to *prevent* racism and industrial accidents. If this amendment is acceptable, then on Dworkin’s view, we should feel a greater responsibility to prevent prejudice against the Deaf community than we should to compensate a person for not having the pleasure of hearing Mozart. This seems right. Nonetheless, this sits uneasily with Dworkin’s view that society owes compensation only for the results of brute bad luck. Getting a cochlear implant is a choice that may allow someone to avoid experiencing societal injustices experienced by deaf people, but it is not right that society is now relieved of the obligation to rectify unjust practices against the Deaf. Indeed, it is ludicrous to suppose that discriminatory practices that target the Deaf are justified if the deaf person in question is a cochlear-refuser. It also does not seem right that the deaf person now has an obligation to get an implant in order to reduce the instances of injustice against her simply because she *can* choose to do so. I likewise would not recommend eliminating sexism by producing only all females or only all males (even if we developed a completely artificial means of procreation, and could produce only females or only males by a less controversial means than abortion). Our obligation is to stop being sexist rather than to stop being a target of sexism, and the obligation rests on the discriminator rather than the discriminated.

In the abstract, the direction of obligation in the case of discrimination is clear. What is less clear is what counts as discrimination, and who holds responsibility in real life situations. For example, the Ontario Human Rights Commission points out that Ontario

law does not require all buildings to have an alarm system with visual features. Most smoke detectors use sound as an alert (although a visual alert is starting to become standard). Ontario law also does not specify who is responsible for covering the cost of providing visual alert systems.⁸

Harshness, Choice, and Responsibility

Dworkin's distinction between brute and option luck is intended to track the distinction between outcomes for which we are responsible and those for which we are not. There are, however, reasons for thinking that choice doesn't always track responsibility. Suppose you receive a live video feed from a serial killer who shows you two captives and tells you that both will die if you do not choose one to die. You choose. While you might *feel* responsible, it seems reasonable to say that the moral culpability belongs to the serial killer alone. We can also generate examples in which there is no morally culpable agent. For example, if a field medic only has enough antibiotics to save one person, and two people need the antibiotics to live, then the field medic has a choice. There may also be no principled reason to choose between the two (neither was reckless, neither is going to go on to cure cancer, etc.). One person will die, and the choice of the medic determines which one dies, but the medic is not morally responsible for that death. It is simply unfortunate. Both of these scenarios are forced choice cases, where the options for producing a preferred outcome are unavailable.

The issue of forced choice also does not always track responsibility either. There are many forced choice scenarios where one does still have responsibilities. Indeed, under certain very harsh conditions that limit choices, responsibility may increase beyond what is normally reasonable. The military, police, firefighters, and survivors in a zombie apocalypse operate under dangerous forced choice scenarios with additional responsibilities. I like the zombie apocalypse scenario for how far it can push our intuitions. Also, some of the extra obligations of the military, police, and firefighters can be attributed to an agreement to take on these obligations, which might include an agreement to maintain a certain level of physical functioning. There is no such agreement in a zombie apocalypse. If it were the case that being deaf would make a person a liability to the survival of the group (and it might not because a deaf person might feel the vibrations of an approaching zombie herd before a hearing person might hear them), then the group might be warranted in

8. <http://www.ohrc.on.ca/en/human-rights-disability-and-accessibility-issues-regarding-visual-fire-alarms-people-who-are-deaf> [last accessed July 13, 2015]

pressuring the acceptance of an implant that would permit hearing, and the deaf person might have an obligation to accept the implant.

Similar considerations could hold for enhancements as well. I find it rather surprising that Liao, Roache, and Sandberg (2012) first argue vigorously for the seriousness of the problem of climate change and the potential that certain enhancements have for mitigating climate change, but then argue that the enhancements should be voluntary. If any situation warrants pressure to enhance, saving the world is it.

Consider a less dire scenario than the end of the world: the mean streets of New York in the 1960s. Given that Spider-Man had great powers, he had great responsibilities. But it isn't clear that Peter Parker had a responsibility to become Spider-Man in the first place. This suggests that the level of extraordinary obligations varies with the level of the severity of the situation. Perhaps in a doomsday scenario, Peter Parker would have an obligation to become Spider-Man.

Earlier we discussed the direction of obligation. Some of the scenarios listed above might be the result of misfortune rather than injustice (the zombie apocalypse might be an "act of God"). We could argue, however, that the climate change case involves injustice. This is a complicated claim involving collective responsibility, obligations to animals, obligations to future persons, lack of knowledge, and so on. But let's suppose we can make the case that some people can be held morally culpable for climate change. Even if we can do this, focusing only on those who created the problem might not provide a sufficient solution to the problem. In the climate change case, Liao, Roache, and Sandberg discuss options for modifying future generations. Obviously future generations did not cause the problem. This suggests that the direction of obligation can be trumped in dire circumstances when even the most strenuous redemptive actions of the culpable would not be sufficient to solve the problem.

A Few Comments on the Real World

Up until now we've discussed some future version of the cochlear implant. The current version is considerably less optimal. Since the cochlear implant involves a medical procedure, it comes with a variety of risks from the medical (infections, nerve damage, etc.) to the economic (e.g., if the manufacturer goes out of business, replacement parts may be difficult to find), and others (e.g., setting off theft detection systems).⁹ Other

9. The FDA has a list of benefits and risks: <http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ImplantsandProsthetics/CochlearImplants/ucm062843.htm> [last accessed August 11, 2015].

considerations include ensuring that the implant does not cause damage to the cochlea that would preclude the use of future, better technologies (Garud and Rappa 1994, 353). It needs to be upgradable. Furthermore, the cochlear implant does not automatically produce hearing, and therapy and education is generally required to support its effectiveness, and the effectiveness varies. The cost of the cochlear runs beyond the cost of the procedure and device itself. It is clear that the decision to get a cochlear implant is a difficult one that could profoundly affect welfare. We might be comfortable with agreeing with Dworkin that it is fine if the lazy person has to sacrifice welfare, but that approach is flippant in this context.

There are social controversies surrounding the cochlear implant as well. In response to social and economic difficulties (including prejudice), many deaf people have participated in the forming of communities, activist societies, and the construction of a Deaf culture. Cultural membership is an important part of identity. To be sure, wine aficionados form groups as well, but this identify-formation generally isn't a response to experiencing prejudice. Comparing the cultural membership of people with expensive tastes to the Deaf culture also seems flippant.

One of the more dramatic social controversies involved a petition to the court to override the decision made by Lee Larson, a deaf single mother of two deaf children. She was encouraged to consider cochlear implants for her children and decided against the procedure. She wanted them to experience Deaf culture for themselves and make their own decisions when they were older. She had to fight in court to have her decision upheld. Many Deaf activists provided vigorous support for Larson. Some of the grounds included protecting the Deaf community as a culture, arguing that deafness is not an illness, and arguing that deafness is not a disability per se, but a disadvantage in a hearing society (Ouellette 2011).

The arguments and motivations behind the Larson court case are interesting. Initially Larson was urged by the school to consider cochlear implants because the school did not have an ASL program. In court the State argued that using spoken language was required for the proper development of the language-processing sections of the brain. This argument was countered by the claim that visual languages also stimulate language-processing abilities (Ouellette 2011, 1248, 1251). The second argument is about the optimal development of the individual. The first argument is about the fit between the individual and society (normalization). It is at this point that we need to think about the direction of obligation.

We also need to think about welfare. Granted, Dworkin is against equalizing welfare, but presumably welfare is not something we want to ignore altogether. There are good reasons to not be too quick to coerce prosthetic use, especially if the focus is on normalization. One fairly obvious reason is that we often make mistakes despite the best of intentions. For example, in Canada, between 1961 and 1962, Thalidomide was available as a treatment for morning sickness for pregnant women. Unfortunately it crossed the placental wall, causing many serious side effects to the foetus (including death). One common side effect is phocomelia, where the limbs are shortened and often shaped like a flipper. The medical advice given to many parents was to outfit their child with prosthetic limbs. This was a focus on normalizing the mode of locomotion, but in many cases it led to a decrease in mobility (level) and an increase in pain (Silvers 114).

When it comes to enhancements, thus far we've talked about them as though we are thinking about some magical medical future. But in all likelihood, it is pretty clear that the various possibilities will not work the same for everyone. Buchanan, Brock, Daniels, and Wikler suggested that we address inequalities at their source, by reducing them rather than compensating for them (2001, 69). As the authors are well aware, however, this may be very difficult to do in practice. Only some will be viable candidates for enhancement treatments, only some will not experience problematic side effects, only some will experience the full benefits, and so on.

Concluding Remarks

My goal in this paper is to explore the extent to which Dworkin's theory has the resources to handle prosthetic and enhancement technologies. It does not. This raises the question of whether we can move towards a theory that can give us better answers while respecting the two sets of intuitions that motivate Dworkin's theory (we may not agree with those intuitions, but that is the subject for a different paper). On the one hand, we have an interest in levelling the playing field. On the other hand, we have an interest in avoiding funding the irresponsible, frivolous, careless, free-rider. We want society to live up to its responsibilities in generating a just society, and we want the individual members to live up to their responsibilities in contributing to society. The problem is that Dworkin's brute/option luck distinction does not fit the bill for allowing us to respect these two sets of intuitions when considering examples of prosthetics and enhancements. In contrast to the lazy and irresponsible, the welfare considerations of those deciding whether or not to get a prosthetic device or an enhancement do not seem at all frivolous.

The brute/option luck distinction does not give us a framework for determining whether individuals are behaving responsibly to the group in part because responsibilities vary with context (level of affluence/poverty of the group, direness of the problem that needs solving, etc.). These contextual matters are often independent of whether or not the individual had a choice to exercise.

In addition, the brute/option luck distinction does not map onto the desire to create a more just, less oppressive society. When an injustice is involved, the direction of obligation matters more than the brute/option luck distinction. Moreover, we want to think about the distinction between backward-looking approaches (compensation for injustices) and proactive approaches (preventing injustices).

Furthermore, real world cases make it clear that it isn't always obvious which course of action will decrease the need for compensatory resources. It also isn't obvious which course of action will increase welfare. If recommendations are put in place with a premature eye to decreasing the need for compensatory resources, this could have the disastrous result of decreasing welfare without actually relieving the economic situation.

We could replace Dworkin's brute/option luck distinction with something else. We could, instead, focus on socially responsible decision making, together with what kind of society we wish to create. Segall, for example, addresses socially responsible decision making by proposing that we "understand "brute luck" as the *outcome of actions (including omissions) that it would have been unreasonable to expect the agent to avoid (or not avoid, in the case of omissions)*" (2010, 20; emphasis in original).¹⁰ Segall emphasizes that this is about what society can reasonably expect, rather than what counts as reasonable behavior on the part of the individual (2010, 20). Obviously the issue of what counts as a reasonable or unreasonable expectation is a complicated discussion, but one worth having nonetheless. I am concerned about having this conversation in the context of resource allocation, however. It will raise again Anderson's (1999) and Denier's (2005) concerns about practicality, intrusiveness, and judgmentalism. The reason I am concerned is because determining allocations is a practical matter that requires that we reach actionable conclusions. Being mistaken is a serious matter in this context. Consequently, we need to have a further conversation about how to proactively reduce the harm potential of our mistakes (and we will make mistakes).

It is wise to think about this from the terms that Anderson sets up. She points out that the egalitarian literature has lost sight of the political agenda of addressing social injustices by focussing more narrowly on correcting for bad brute luck (1999, 288). If

10. Thanks to Martin Gunderson for bringing this work to my attention.

we focus on correcting for brute bad luck, our attention gets directed to the question of whether the luck is brute or option. If we follow Segall's redefinition of brute luck, then we'll focus on whether society can reasonably expect a certain choice. If, instead, we focus on addressing social injustices, one agenda that comes to the fore is that we need to pay attention to how society generates choices for individuals. This then influences what society can reasonably expect from individuals. Indeed, instead of thinking only about how to place economic pressure on individuals to encourage them to make responsible decisions, we could also pay attention to putting economic pressure on society to make it economically rational to reduce systems of oppression and to generate reasonable choices for its citizens.

References

- Anderson, Elizabeth S. 1999. "What is the point of equality?" *Ethics* 109 (2): 287-337.
- Arneson, Richard J. 1982. "The principle of fairness and free-rider problems." *Ethics* 92 (4): 616-633.
- Buchanan, Allen, Dan W Brock, Norman Daniels, and Daniel Wikler. 2001. *From chance to choice: genetics and justice*. Cambridge: Cambridge University Press.
- Burley, Justine (ed). 2004. *Dworkin and his critics: with replies by Dworkin*. Oxford: Blackwell Publishing Ltd.
- Crossley, Mary. 2003. "Reasonable accommodation as part and parcel of the antidiscrimination project." *Rutgers LJ* 35: 861.
- Denier, Yvonne. 2005. "On personal responsibility and the human right to healthcare." *Cambridge Quarterly of Healthcare Ethics* 14 (02): 224-234.
- Dworkin, Ronald. 1981. "What is equality? Part 2: equality of resources." *Philosophy & Public Affairs* 10 (4): 283-345.
- Dworkin, Ronald. 2000. *Sovereign virtue: the theory and practice of equality*. Cambridge, Massachusetts: Harvard University Press.
- Dworkin, Ronald. 2002. "Sovereign virtue revisited." *Ethics* 113 (1): 106-143.
- Dworkin, Ronald. 2003. "Equality, luck and hierarchy." *Philosophy & Public Affairs* 31 (2): 190-198.
- Fleurbaey, Marc. 2005. "Freedom with forgiveness." *Politics, Philosophy & Economics* 4 (1): 29-67.
- Fleurbaey, Marc. 2008. *Fairness, responsibility, and welfare*. New York: Oxford University Press.
- Garud, Raghu, and Michael A Rappa. 1994. "A socio-cognitive model of technology evolution: The case of cochlear implants." *Organization Science* 5 (3): 344-362.
- Griffin, James. 1986. *Well-Being: its meaning, measurement, and moral importance*. Oxford: Oxford University Press.
- Liao, Matthew, Anders Sandberg, and Rebecca Roache. 2012. "Human engineering and climate change." *Ethics, Policy & Environment* 15 (2): 206-221.
- Ouellette, Alicia. 2010. "Hearing the deaf: cochlear implants, the deaf community, and bioethical analysis." *Val. UL Rev.* 45:1247.
- Rawls, John. 1971. *A theory of justice*. Cambridge: Harvard University Press.

- Ryan, Cheyney C. 1980. "The normative concept of coercion." *Mind* 89 (356): 481-498.
- Segall, Shlomi. 2010. *Health, luck, and justice*. Princeton University Press.
- Silvers, Anita. 1998. "A fatal attraction to normalizing: Treating disabilities as deviations from species-typical functioning." In *Enhancing Human Traits: conceptual complexities and ethical implications*, Eric Parens (ed.). Washington DC: Georgetown University Press: 95-123.
- Wertheimer, Alan. 1987. *Coercion*. Princeton: Princeton University Press.
- Wikler, Daniel. 2004. "Personal and social responsibility for health." In *Public health, ethics, and equity*, edited by Anand, Sudhir, Fabienne Peter and Amartya Sen, 109-134. Oxford: Oxford University Press.

Journal of Cognition and Neuroethics

Reversibility and Deep Brain Stimulation

Jennifer Mundale

University of Central Florida

Biography

Jennifer Mundale is an Associate Professor of Philosophy and the Cognitive Sciences at the University of Central Florida. Her research interests include Cognitive Science, Philosophy of Neuroscience, and Philosophy of Psychology, particularly Clinical Psychology.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). January, 2016. Volume 3, Issue 4.

Citation

Mundale, Jennifer. 2016. "Reversibility and Deep Brain Stimulation." *Journal of Cognition and Neuroethics* 3 (4): 97–111.

Reversibility and Deep Brain Stimulation

Jennifer Mundale

Abstract

Reversibility is a much-touted advantage of Deep Brain Stimulation (DBS) and related neuromodulatory procedures. In the treatment of both motor and psychological disorders, earlier surgical procedures aimed at the permanent ablation of specific brain areas, but DBS, in contrast, does not deliberately seek to destroy brain tissue. For this and other reasons I discuss, DBS is widely described as “reversible”, and it is this claim of reversibility that is the focus of my essay. I argue that, not only is there no common agreement about what “reversibility” means, there are important respects in which the claim is false, and others in which it is misleading.

Keywords

DBS, Deep Brain Stimulation, Psychosurgery, Reversibility, Movement disorders, Neurophilosophy, Neuroethics

...whilst we think, our brain changes, and that, like the aurora borealis,
its whole internal equilibrium shifts with every pulse of change.
—William James, *Principles of Psychology*

Introduction

Reversibility is a much-touted advantage of Deep Brain Stimulation (DBS) and related neuromodulatory procedures. DBS is best known in the treatment of Parkinson Disease and other movement disorders, but recently, researchers have begun to treat a wider array of disorders with DBS, including, more controversially, psychiatric ones. In the treatment of both motor and psychological disorders, earlier surgical procedures aimed at the permanent ablation of various, targeted brain areas, so the appeal of a neuromodulatory alternative – one which does not deliberately seek to destroy brain tissue – is perhaps obvious. For this and other reasons explored below, DBS is commonly said to be *reversible*.

It is this claim of *reversibility* that is the focus of this critique. Importantly, it is not intended to be a general commentary about the overall therapeutic value of deep brain stimulation nor whether it is an appropriate course of treatment in any given case. These are questions for medical professionals, and highly specialized ones at that, neither of which I am. It is also important to note that some of the disorders and diseases treatable

with DBS can be severely painful, disabling, and otherwise intractable; such cases may warrant the risk of irreversible changes or damage to the patient's brain or overall psychological condition. Similarly, Parkinson Disease and some of the other disorders treated with DBS are themselves sufficiently damaging such that the risk not only of permanent changes from the treatment, but the possibility of damage from the disease itself, has to be assessed in the context of risks vs. possible gains.

To the extent, however, that contemplation of DBS treatment requires an informed decision on the part of the patient, a more careful analysis of its claims of "reversibility" would seem to be an obvious and integral consideration in the decision process. Unfortunately, the answers to the following questions are anything but clear: What specific aspects or components of DBS are reversible, and what, exactly does "reversible" amount to, in this context? In the context of DBS, *reversibility* enjoys neither universal applicability nor unambiguous understanding.

In what follows I consider various notions of reversibility, examine sample cases of claimed reversibility, and argue that the reversibility of DBS should by no means be taken for granted. Furthermore, the medical uncertainty of the reversibility of both desirable as well as unintended and/or undesirable effects points us toward important and unresolved ethical difficulties concerning DBS. Ongoing controversies concerning its side effects, possible alteration of patients' personalities, mood, cognitive abilities, identities, and sense of autonomy have been widely addressed in the literature, but less attention has been given to the claim that DBS can somehow, reassuringly, all be reversed, and the former controversies only heighten the importance of acknowledging questions of reversibility. Furthermore, as DBS is constantly expanding to new medical applications and likely has not yet realized its full scope or potential, a more thorough appreciation of these questions becomes more significant.

Applications of Deep Brain Stimulation

Although various experimental and diverse methods of electrical brain stimulation can be documented historically, modern DBS is commonly taken to have been in use since 1987 (Ineichen et al. 2014) with the landmark advances of Benabid, et al., in the treatment of Parkinsonism (Benabid et al. 1987). Since that time, over 100,000 patients worldwide have undergone the procedure. This number is rapidly growing, as is the associated medical research and literature. Since 2000, over 8,000 medical journal articles have been published on DBS since 2000, and over 1,000 have appeared just in the past year (as determined by EBSCO search). DBS is still best known and most commonly used

for the treatment of Parkinson's Disease and other movement disorders. These are the earliest modern applications of DBS, and so far, the best understood. They also constitute the only category of applications for which data can be collected about the long-term effects of DBS in a large enough number of patients to be of statistical significance. The ubiquitous claims of reversibility, therefore, may be tacitly resting on these standard sorts of cases which, because of their historical precedence, provide most of the existing clinical data.

In recent years, however, the therapeutic application of DBS has gradually crept into applications involving disorders which, themselves, are not always thoroughly understood, either with respect to the neural mechanisms underlying their manifestation and control, or the diagnostic criteria that identify them. Now, a small but growing number of cases have been treated with DBS for such things as chronic pain, headaches, obsessive compulsive disorder, Tourette Syndrome, severe depression, bipolar disorder and morbid obesity (Dormont et al. 2009). In 2007, DBS was even used in five subjects with writer's cramp that had not responded to other treatments (Fukaya et al. 2007). There is a particularly strong research trend to apply DBS to various psychiatric disorders.

Related to the growing application to psychiatric disorders is a related trend, though one that is still largely still prospective, and that is the treatment of criminal and sociopathic patients with DBS. Perhaps most noteworthy, Italian neurosurgeon Sergio Canavero, has strongly advocated extending various forms of psychosurgery, including DBS and other neuromodulatory procedures to those who engage in criminal or violent behavior and/or who suffer from drug addiction. His recommendations are premised on several controversial assumptions, including that "Free will is a mere illusion", that "Psychopathic behavior is a purely biological epiphenomenon and can be induced", and, with respect to criminal treatment, "The goal is redirecting the action course of the criminal behavior by 'rewriting' the original priming signal to commit an antisocial act" (Canavero 2014). As Canavero is also famous for proclaiming that he will perform the first human head transplant (Canavero 2013), and has attracted criticism on ethical grounds for this (e.g., Kaplan, 2015; Čartolovni and Spagnolo 2015), one should be cautious about regarding his views as representative of mainstream neurology. Nonetheless, the cases he cites as pointing the way toward the treatment of criminal insanity are real, and judging from those and other recent research, there are genuine moves in this direction.

In 2010, for example, a Tulane research team used DBS to treat a nineteen year old woman diagnosed with intermittent explosive disorder. This patient also had been diagnosed with moderate mental retardation and bipolar disorder, and had been treated with various psychotropic medications. She was reported to have tolerated the

surgical procedure itself well, but the period of adjusting the neuromodulatory settings took approximately a year, during which the patient was clinically depressed, showed symptoms of obsessive-compulsive disorder, overdosed on her medication and had to be committed to a psychiatric ward for 3 months. After this adjustment period, the authors report that, “the goals of attenuating aggressive impulses and providing the patient with control over her emotions and violent outbursts were achieved. A significant improvement in the quality of life of both the patient and her family was seen almost immediately upon determining the proper settings of her stimulator. ... we found that there was a fine line between achieving control of symptoms and producing some depression as well as obsessive-compulsive disorder symptoms” (Maley et al. 2010).

Such applications raise the usual sorts of ethical concerns alluded to earlier, including informed consent, alterations in patient identity, possible restrictions of autonomy, and the need to consider the implications of such surgeries in light of a general, precautionary principle. As noted earlier, these ethical issues have been more widely discussed, and while not the direct concern of this paper, they do make questions about the reversibility of such procedures more compelling.

Claims of Reversibility

One of the chief advantages *claimed* for DBS is that it is *reversible*. Perhaps in the public mind the emphasis of this feature is meant to counter the inevitable specter of past psychosurgical abuses. As noted above, DBS aims toward electrical *stimulation* (neuromodulation), rather than the *irreversible ablation*, of targeted brain tissues. This is in marked contrast with, for example, Walter Freeman’s “icepick lobotomies” (in which real icepicks were driven by a hammer through the orbital roof into the frontal lobe), or the crude, early forms of electroconvulsive therapy which were performed without restraints or anaesthetics and produced such violent convulsions that patients’ bones and vertebrae often fractured.

Reversibility is also a widely publicized claim in the DBS literature; one finds it in nearly every medical overview, on many patient-oriented websites, and in much of the professional and scientific research on DBS. While DBS has been the focus of ethical concerns, those concerns have largely addressed issues such as safety, patient selection, informed consent, negative psychological side-effects, patient autonomy, personal identity, and its experimental use for psychiatric, and other non-movement related disorders (see, for example, Synofzik, 2015, Clausen, 2010, Schermer 2011). Comparatively little critical attention has been given to the claim of reversibility itself.

In the table below are representative samples of claims of reversibility drawn from sources intended for medical professionals, and also sources meant for patients (the claims of reversibility are unmanageably numerous in the medical and patient literature so this only attempts a small, but reasonable sampling). An examination of these claims shows that, in each case, the nature of the claimed reversibility is subtly different.

The first case is typical for its vagueness; it is the “therapy”, non-specifically stated, that is said to be reversible. The second case is more specific, and claims that the “modulation of faulty neural circuits” is reversible, suggesting that those faulty circuits can be returned to their pre-modulated state, or the state they were in before the stimulation was applied. The third claim involves reversing the “functional ablation.” In this context, a functional ablation means that the activity of the area targeted by the stimulation is suppressed, so the reversibility claim here suggests that the area will return to a state of pre-stimulation functioning when the suppressing stimulation is stopped. This “functional ablation” stands in sharp contrast with older surgeries that involved the intentional, permanent, physical destruction of brain tissue, and, as noted earlier, it is a contrast many researchers seem particularly concerned to promote.

The fourth claim is similar to the one before, but its target audience is current or prospective DBS patients so may be less technical. What is claimed is a reversible alteration to abnormally functioning brain tissue. The last case, like the former, is meant for the general public consumption. It vaguely, but quite boldly proclaims that “The procedure is entirely reversible, usually with minimal damage to any brain tissue.” This seems somewhat contradictory, in that one might wonder how something could be entirely reversible if it causes *any* damage to brain tissue, minimal or otherwise, but perhaps the claim involved here is that the damage to the brain tissue is also reversible. The trouble, of course, is that it is not at all clear what exactly is included within the scope of the term “procedure”, so the claim of reversibility is vacuous, if not falsely reassuring. Furthermore, it is a poor basis on which to build an informed patient and medical community.

Claim	Source
1) "Since 1999, the success of deep brain stimulation, a new reversible and adaptable therapy devised for the treatment of Parkinson's disease, has offered the hope of new forms of treatment for patients with severe psychiatric disorders like OCD."	(Medically specialized text) Lévêque, M. 2013. <i>Psychosurgery: New Techniques for Brain Disorders</i> . Paris: Springer-Verlag, p. 22.
2) "DBS is appealing because it provides precisely targeted, adjustable, and reversible modulation of faulty neural circuits occurring in a variety of brain disorders."	(Medically specialized text) Marks, WJ. 2015. "Preface." In <i>Deep Brain Stimulation Management</i> , 2 nd edition, edited by WJ Marks, xi. Cambridge: Cambridge University Press.
3) "Such surgical experiences as reversible functional ablation have been applied to deep brain stimulation (DBS) of thalamus to date, and the most promising surgical target for intractable tremor of PD is the nucleus ventrointermedius (Vim) of the thalamus.	(Medically specialized text) Miyagi, Y. 2015. "Thalamic Stimulation for Parkinson's Disease: Clinical Studies on DBS." In <i>Deep Brain Stimulation for Neurological Disorders: Theoretical Background and Clinical Application</i> , edited by T. Itakura, 104. Dordrecht, Netherlands: Springer.
4) "Unlike older lesioning procedures or gamma knife radiosurgery, DBS does not destroy brain tissue. Instead, it reversibly alters the abnormal function of the brain tissue in the region of the stimulating electrode."	(Patient information site) University of Pittsburgh, Neurological Surgery, website for patients: http://www.neurosurgery.pitt.edu/centers-excellence/epilepsy-and-movement-disorders-program/deep-brain-stimulation-movement-disorders
5) "DBS differs from pallidotomy, thalamotomy, and subthalamotomy in that it does not permanently destroy brain tissue. The procedure is entirely reversible, usually with minimal damage to any brain tissue."	(Patient information site) National Parkinson Foundation, patient literature: Okun, MS., and PR. Zeilman. 2014. <i>Parkinson's Disease: Guide to Deep Brain Stimulation Therapy</i> . 2 nd ed. National Parkinson Foundation, 5. http://www.parkinson.org/sites/default/files/Guide_to_DBStimulation_Therapy.pdf

The next section attempts to further evaluate different aspects of DBS reversibility, but it is in the nature of the problem that perfect clarity is not to be had. In some cases such claims are demonstrably false, in some others likely true, and in some cases, the state of current scientific understanding is not sufficiently developed to know whether such a claim is true. To best examine these questions, however, it is useful to divide the question into two broad categories; reversibility of the implantation procedure, and reversibility of the effects.

Reversibility and the Implantation Procedure

A current DBS research team puts things succinctly when they say: “DBS has advantages in reversibility and adjustment, but disadvantages in device implantation” (Nishibayashi and Itakura 2015). With respect to the implantation procedure, it is useful to note that the stimulation device consists of three distinct parts: the battery powered pulse generator, or main unit, which is implanted under the skin (usually near the collarbone, though other locations are occasionally used), one or two leads (generally with 4 contacts on each lead) that are inserted into the targeted brain tissue, and a wire that connects the leads to the pulse generator. While procedures for lead placement vary and are in constant development, a common technique involves placement of leads as guided by a rigid stereotactic frame, which is attached to the outside of the patient’s head while the patient’s brain is imaged by an MRI prior to the to identify the target structures in the brain relative to the frame. In this procedure, the patient is awake during the implantation of the leads (but not in pain), and can give responses to the surgeon to help determine correct lead placement. Recently, results from a variation on this procedure heralded greater accuracy in lead placement, and, in addition to a pre-operative MRI, uses CT scans during the procedure itself, which is conducted on sleeping, rather than alert patients (see Burchiel et al. 2013).

Reports vary about the effects of the lead insertion itself. While many report that this part of the procedure has no lingering effects, neurologist Paul Foley, in a 2015 paper, for example, argues that the mere insertion of the electrodes, can cause irreparable tissue damage, “the long-term consequences of which are unknown” (2015 565). Since detailed, microscopic sectioning and examination of the tissue around the lead cannot happen until post mortem, human data is limited. As with much DBS patient data, reported results vary considerably depending on such factors as the tissues targeted, the disorders being treated, the condition of the patient, duration of the treatment, the clinic at which the surgery is performed, the surgical team, and several other variables.

Most patients, however, show at least some degree of microgliosis; a complex process of localized changes in the glial cells in response to injury. This, too, varies by degree (see, for example, Sun, D.A. et al. 2008).

The issue is more complicated than first appears not only because of the variability and sparseness of results, but because the so-called “microlesioning” of the targeted tissue with the lead itself sometimes provides a noticeable, though temporary, therapeutic contribution, sometimes referred to as the “honeymoon period” (Kumar and Johnson 2011). While techniques are improving every year, in a 2008 text on neural implants generally (including DBS), one researcher summarized the situation as follows:

It is critical to understand the nature and mechanisms of the tissue response to the implantation, residence, and in the case of stimulation, activation of electrodes in the CNS. These devices are rapidly becoming more widespread, smaller, and more dense. Unfortunately, there remains a lack of first principles understanding of the mechanisms of neuronal injury. Thus, the issue of damaging versus nondamaging neural interfaces has been and will continue to be addressed in a purely empirical manner. Analysis of postmortem human and animal tissue has shown that there is neuronal loss around chronically implanted electrodes and a high density of astrocytes, microglia, and vasculature around the electrode. Loss of neurons around the electrode may affect how well the neural prosthesis functions, especially as devices move toward smaller arrays of electrodes that use microstimulation... (Grill, 2008)

The author further notes that the higher voltages associated with OCD, for example, carry an increased risk of tissue damage.

There is also considerable variation in the number of surgical or device-related complications from one institution to another. A 2006, six year retrospective from Newcastle General Hospital, for example, reports a complication rate of 30%:

During the study period, a total of 60 patients underwent 96 procedures for implantation of unilateral or bilateral DBS electrodes. The mean follow-up period was 43.7 months (range 6-78 months) from the time of the first procedure. No patients were lost to follow-up or died. Eighteen patients (30%) developed 28 adverse events, requiring 28 electrodes to be replaced. Seven patients developed two adverse

events and two patients developed three adverse events. The rate of adverse events per electrode-year was 8%. (Paluzzi, et al., 2006). On the other hand, a ten-year retrospective of patients treated at a Swedish hospital, published in 2005, reported only a 15% hardware-related complication rate, and others still lower. (Blomstedt et al. 2005)

In addition to cellular changes surrounding the electrode, there is, as with any surgery, risk of brain hemorrhage. Estimates vary, but the risk during removal – of concern to the question of reversibility – is greater than during the original implantation. A 2005 retrospective study involving 78 DBS lead removals showed a risk of intracranial hemorrhage of 12.8% per lead for removal, but only 2.0% per lead for implantation (at the authors' clinic). It is important to also point out that, according to the authors, "all hemorrhages were asymptomatic" (Liu 2012).

As mentioned previously, the therapeutic value or medical advisability is not the focus here. The concern, rather, is with the frequently touted claim of reversibility, and on this point, the picture appears to be mixed with respect to the implantation of the leads. If reversibility means a return to a state or condition that existed prior to the procedure, the implantation of the leads appears pretty plainly to rule this out. The physical, irreversible changes, one may argue, are of minimal importance, but in some cases, those changes result in permanent, functional changes as well. Moreover, the extent of such changes are variable and unpredictable.

Reversibility and the Effects of DBS

The simplest, and most direct sense in which DBS is commonly taken to be reversible is this: when the current that provides the electrical stimulation is stopped, the effects of the current also stop (see, for example, Yu and Neimat 2008; Machado et al. 2012). An examination of the literature, however, shows that this claim, that the effects stop when the current stops, is at best, problematic. Even a cautious defender of DBS points out, that, "While the stimulation might be reversible, it remains an open question to what extent the effects of the stimulation are indeed reversible, and it is the effects which are morally relevant" (Pacholczyk 2015, 641).

For example, with respect to the specific context of DBS's positive effects, one curiously finds discussions involving the *retention of clinical benefit*, rather than reversibility (or irreversibility). Some studies indicate that the therapeutic value of DBS can linger for an indefinite period of time after the neurostimulating current is turned off. One of the earliest reports documenting this longterm effect appeared in 2007 (Hebb

et al.) and followed the clinical course a single patient. In a remarkably frank opening statement, the authors write:

Deep Brain Stimulation (DBS) is an effective treatment option for various movement disorders and is being investigated for use in chronic pain, epilepsy, and select neuropsychiatric conditions. This growing list of indications has superseded our knowledge of either the short- or long-term physiological effects of high frequency stimulation (HFS) in the human brain. Although reversibility is a touted hallmark of DBS, other findings in these patients may allude to more long-term changes taking place in the brain as a result of chronic exposure to HFS. (1958-9)

In the particular patient they followed, the patient was able to stop the stimulation after five years, and was monitored for a year following its cessation. In this patient's case, they speculate as follows: "It is probable that therapy-induced plasticity within the involved circuits contributed to these effects and further study is needed to discern the physiological sequelae of long-term DBS" (1961). Such study is hampered by the understandable and necessary limitations in studying living subjects. As they explain, "Unfortunately, such studies are difficult or impossible to perform in vivo and there are currently no direct methods of evaluating LTP or LTD [long term potentiation and depression] in the living, in situ human brain" (1961).

Following this 2007 study, some researchers have sought to understand, develop and shape the long term effects of DBS to expand its therapeutic potential. One such article by Ruge, et al., follows another patient with dystonia, in whom "there was no change in average physiological or clinical status when deep brain stimulation was turned off for 2 days, suggesting that deep brain stimulation had produced long-term neural reorganization in the motor system" (2011, 2106-7). Even more recently, a group with the same lead author reported in the *Journal of the Neurological Sciences* that, "during early stages of therapy, dystonia patients often revert back quickly to their pre-operative state when DBS is switched OFF whilst after several years of DBS the beneficial clinical effect in some patients can be retained for long periods" (Ruge et al. 2014, 197-199).

In the kinds of cases described in this section, reversibility may not be desirable since they point to longterm therapeutic value for the patient. Nonetheless, the longterm physiological and changes this research suggests casts further doubt on the idea that DBS can be reversed by simply shutting it off.

Conclusion

I have argued that the claim DBS is reversible is fraught with many difficulties, and at very least, should not be taken for granted. One might claim that the many empirical unknowns and uncertainties in such a new therapy make truly informed consent an impossible epistemological standard. One might also point to the conceptual difficulties of reversibility itself and take the position that *no* mental or physical changes are truly or strictly reversible: *thinking* causes irreversible changes, as William James pointed out over a century ago. These uncertainties and difficulties, however, only underscore the need for greater transparency and candor. The goal of informed patient consent would be better served by replacing the near ubiquitous, over-simplified claims for the reversibility of DBS with a more accurate and better contextualized explanation of changes that may persist indefinitely, and a candid admission of the many uncertainties that accompany them. The extent of these deficiencies points us toward important and unresolved ethical challenges concerning DBS.

Bibliography

- Benabid, AL., P. Pollak, A. Louveau, S. Henry, and de Rougemont. 1987. "Combined (thalamotomy and stimulation) stereotactic surgery of the VIM thalamic nucleus for bilateral Parkinson disease." *Applied Neurophysiology* 50 (1-6): 344-6.
- Blomstedt, P., and MI Hariz. 2005. "Hardware-related complications of deep brain stimulation: a ten year experience." *Acta Neurochirurgica* 147 (10): 1061-4.
- Burchiel, KJ., S. McCartney, A. Lee, and AM. Raslan. 2013. "Accuracy of deep brain stimulation electrode placement using intraoperative computed tomography without microelectrode recording." *Journal of Neurosurgery* 119 (2): 301-306.
- Canavero, S. 2013. "HEAVEN: The head anastomosis venture Project outline for the first human head transplantation with spinal linkage (GEMINI)." *Surgical Neurology International* 13 (4) (Suppl 1):S335-42.
- Canavero S. 2014. "Criminal minds: neuromodulation of the psychopathic brain." *Front. Hum. Neurosci* 8 (124).
- Čartolovni, A., and AG. Spagnolo. 2015. "Ethical considerations regarding head transplantation." *Surgical Neurology International* 6 (103).
- Clausen, J. 2010. "Ethical brain stimulation - neuroethics of deep brain stimulation in research and clinical practice." *European Journal of Neuroscience* 32 (7): 1152-1162.
- Dormont, D., D. Seidenwurm, D. Galanaud, P. Cornu, J. Yelnik, and E. Bardinet. 2009. "Neuroimaging and Deep Brain Stimulation." *American Journal of Neuroradiology* 2010 31: 15-23.
- Foley, P. 2015. "Deep brain Stimulation for Parkinson's Disease: Historical and neuroethical aspects." In *Handbook of Neuroethics*, edited by J. Clausen and N. Levy. Dordrecht, Netherlands: Springer.
- Fukaya, C., Y. Katayama, T. Kano, T. Nagaoka, K. Kobayashi, H. Oshima, and T. Yamamoto. 2007. "Thalamic deep brain stimulation for writer's cramp." *Journal of Neurosurgery* 107 (5): 977-982.
- Grill, WM. 2008. "Signal Considerations for Chronically Implanted Electrodes for Brain Interfacing." In *Indwelling Neural Implants: Strategies for Contending with the In Vivo Environment*, edited by W. M. Reichert. Boca Raton, Florida: CRC Press/Taylor and Francis.

- Hebb, MO., P. Chiasson, AE. Lang, RM. Brownstone, and I. Mendez. 2007. "Sustained relief of dystonia following cessation of deep brain stimulation." *Movement Disorders* 22 (13): 1958-62.
- Ineichen, C., W. Glannon, Y. Temel, CR. Baumann, and O. Sürücü. 2014. "A critical reflection on the technological development of deep brain stimulation (DBS)." *Frontiers in Human Neuroscience* 8 (730).
- Kumar, R., and L. Johnson. 2011. "Managing Parkinson's disease patients treated with deep brain stimulation." In *Deep Brain Stimulation Management*, edited by William J. Marks, Jr, 62-82. Cambridge: Cambridge University Press.
- Lévêque, M. 2013. *Psychosurgery: New Techniques for Brain Disorders*. Paris: Springer-Verlag.
- Liu, JK., H. Soliman, A. Machado, M. Deogaonkar, and AR. Rezai. 2012. "Intracranial hemorrhage after removal of deep brain stimulation electrodes." *Journal of Neurosurgery* 116 (3): 525-8.
- Machado, AG., M. Deogaonkar, and S. Cooper. 2012. "Deep brain stimulation for movement disorders: patient selection and technical options." *Cleveland Clinic Journal of Medicine* Jul 79 (Suppl 2): S19-24.
- Maley, JH., JE. Alvernia, EP Valle, and D. Richardson. 2010. "Deep brain stimulation of the orbitofrontal projections for the treatment of intermittent explosive disorder." *Neurosurgical Focus* 29 (2).
- Marks, WJ. 2015. "Preface." In *Deep Brain Stimulation Management*, 2nd edition, edited by WJ Marks, xi. Cambridge: Cambridge University Press.
- Miyagi, Y. 2015. "Thalamic Stimulation for Parkinson's Disease: Clinical Studies on DBS." In *Deep Brain Stimulation for Neurological Disorders: Theoretical Background and Clinical Application*, edited by T. Itakura, 104-120. Dordrecht, Netherlands: Springer.
- Nishibayashi, H., and T. Itakura. 2015. "Surgical Technique of Brain Stimulation." In *Deep Brain Stimulation for Neurological Disorders: Theoretical Background and Clinical Application*, edited by T. Itakura, 69. Dordrecht, Netherlands: Springer.
- Okun, MS., and PR. Zeilman. 2014. *Parkinson's Disease: Guide to Deep Brain Stimulation Therapy*, 2nd ed., National Parkinson Foundation.

- Pacholczyk, A. 2015. "Ethical objections to deep brain stimulation for neuropsychiatric disorders and enhancement: A critical review." In *Handbook of Neuroethics*, edited by J. Clausen and N. Levy. Dordrecht, Netherlands: Springer.
- Ruge, D., L. Cif, P. Limousin, V. Gonzalez, X. Vasques, Ml. Hariz, P. Coubes, and JC. Rothwell. 2011. "Shaping reversibility? Long-term deep brain stimulation in dystonic: the relationship between effects on electrophysiology and clinical symptoms." *Brain: A Journal of Neurology* 134: 2106-2115.
- Ruge, D., L. Cif, P. Limousin, V. Gonzalez, X. Vasques, P. Coubes, and JC Rothwell. 2014. "Longterm deep brain stimulation withdrawal: Clinical stability despite electrophysiological stability." *Journal of the Neurological Sciences* 342 (1): 197-199.
- Schermer, M. 2011. "Ethical Issues in Deep Brain Stimulation." *Frontiers in Integrative Neuroscience* 5 (17).
- Sun, DA., H. Yu, J. Spooner, AD. Tatsas, T. Davis, TW. Abel, C. Kao, PE. Konrad. 2008. "Postmortem analysis following 71 months of deep brain stimulation of the subthalamic nucleus for Parkinson disease." *Journal of Neurosurgery* 109 (2): 325-329.
- Synofzik, M. 2015. "Ethical implications of brain stimulation." In *Handbook of Neuroethics*, edited by J. Clausen and N. Levy. Dordrecht, Netherlands: Springer.
- Paluzzi, A., A. Belli, P. Bain, X. Liu, and TM Aziz. 2006. "Operative and hardware complications of deep brain stimulation for movement disorders." *British Journal of Neurosurgery*. 20 (5): 290-5.
- Yu, H., and JS. Neimat. 2008. "The treatment of movement disorders by deep brain stimulation. *Neurotherapeutics*." 5 (1): 26-36.

Journal of Cognition and Neuroethics

A Subjectivist Solution to the Problem of Harm in Genetic Enhancement

Sruthi Rothenfluch
University of Portland

Biography

Sruthi Rothenfluch is a visiting instructor of philosophy at the University of Portland in Oregon. She completed her doctorate at the University of Nebraska-Lincoln in 2011, working primarily within epistemology. Her current research centers on the ethics of involuntary medical intervention as this applies to pre-natal genetic enhancement and psychiatric treatment. Her previous work within contextualism and virtue epistemology, which addressed differences between expert and lay deliberation and judgment, raised more practical questions about the content of expert claims and whether they ought to be invariably accepted. This led her to examine conditions under which treatments may be permissible or even obligatory in the absence of consent.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). January, 2016. Volume 3, Issue 4.

Citation

Rothenfluch, Sruthi. 2016. "A Subjectivist Solution to the Problem of Harm in Genetic Enhancement." *Journal of Cognition and Neuroethics* 3 (4): 113–130.

A Subjectivist Solution to the Problem of Harm in Genetic Enhancement

Sruthi Rothenfluch

Abstract

Some have recently argued that parents are morally obligated, under certain circumstances, to use pre-natal genetic intervention as a means of enhancement. Despite aiming to benefit the child, such intervention may produce serious and irreparable harm. In these cases, parents seem to have an obligation not to intervene, as such efforts make the child worse off. Julian Savulesu has argued that while harm raises doubts about the acceptability of genetic enhancement, genetic selection remains an obligation. This claim, however, rests on an indefensible privileging of personal over impersonal harm. I propose instead that we reframe the debate as stemming from fundamentally different views about parental obligation. The objection from harm rests on an objectivist conception, according to which obligation is determined by all relevant facts, including unpredictable harm. Proponents of genetic enhancement, however, operate within subjectivist assumptions about obligation, according to which moral requirements are determined by reasons that are epistemically accessible to the relevant agents. I will argue here that because subjectivism offers a more reasonable conception of parental obligation, such unforeseeable harm does not remove a parent's obligation to enhance.

Keywords

Genetic Enhancement, Genetic Selection, Subjectivism, Harm, Procreative Beneficence

I. Introduction

While the ethical status of pre-natal genetic enhancements is far from settled, both supporters and opponents of such technology view the child's interests and safety as central to the debate. Opponents often cite harm to the child in terms of loss of autonomy and distortion of the parent-child relationship as reasons against enhancement (Sandel 2007) while Julian Savulescu and other proponents emphasize ways in which enhancement will improve the quality of a child's life. Rather than unrestricted approval of all such interventions, the latter generally acknowledge the permissibility of enhancement under certain specified conditions, in particular, those cases in which we have strong evidence that more good will come about as a result of enhancement than without it.¹ These goods have been fleshed out in positive and negative ways, citing both the best interests of the recipient, and the absence of conflicting considerations

1. Brock 2009; Buchanan et al 2000; Robertson 2001

such as parents' impaired ability to care for existing children, and harmful impacts on the wider society. Savulescu has argued that under certain conditions, prenatal genetic interventions are not only permissible, but *required*, on par with providing nutrition, education and shelter.² While this might reek of a move towards eugenics, Savulescu and others, as will be shown below, do not endorse a narrowly defined conception of a good life, but instead remain neutral on this question. Despite its merits, the position appears susceptible to a certain kind of charge: suppose that as a result genetic enhancement, you produce a child who is predisposed to asthma, or worse, cancer. In such cases, it appears that the child's life turned out worse due to intervention. Savulescu (2001, 2006) appeals to a *person-affecting* sense of harm to argue that while genetic enhancement may harm the child, selection will not. This is because *selecting* an embryo with genes predisposed to serious illness does not make that child any worse. If parents had selected an embryo without that genetic structure, it would have been a different child. Therefore, since the child is not worse than she would have been, no harm has been done to that child. Such a view of harm is overly restrictive in that it precludes, (or, at best, fails to appreciate the severity of) genuine instances of harm. Instead of denying that intervention causes harm, I propose that the best way to understand and respond to the objection is to expose its objectivist underpinnings. Harm that was not and could not have been reasonably predicted can count against an act only if all relevant facts determine its moral status. Because parental obligations are best determined subjectively, that is, according to facts parents can access, the obligation to enhance can be sustained in such cases.

I begin in the second section by presenting Savulescu's extensive defense of the moral obligation to enhance, supporting and supplementing his view at times with others' whose positions strongly resemble his own. As I largely agree with Savulescu, I will go some way towards clarifying and defending his position. Next, I will present a case that is representative of the worry described above and explain Savulescu's response that selection is immune from this objection. I will, however, contend that his strategy fails to adequately address the problem. In the final section, I will present my own proposal for re-assessing the debate as one instance of the larger subjectivist-objectivist divide within metaethics. I will argue that because parental obligations are determined subjectively, the objection from harm cannot displace our obligation to use ethical genetic enhancements. I will end by addressing some apparently problematic cases for my view.

2. Savulescu and Kahane 2009; Savulescu 2010, 2009, 2001; Harris 2010

II. Principle of Procreative Beneficence (PB)

Savulescu and Kahane (2009) endorse what they call the 'Principle of Procreative Beneficence' (PB), one version of which is given here:

If couples (or single reproducers) have decided to have a child, and selection is possible, then they have a significant moral reason to select the child, of the possible children they could have, whose life can be expected, in light of the relevant available information, to go best or at least not worse than any of the others. (274)³

The basic principle here is not difficult to understand or support, insofar as it encourages parents to aim at having the best child possible, something that parents routinely do outside of genetic intervention. Savulescu (2009b) has argued that because genetic enhancements are relevantly similar to environmental enhancements in their effects and expectations, the two are morally equivalent. I will not rehearse his arguments here because my chief concern is whether an obligation to enhance can be sustained against the objection from harm. There are some aspects, however, that call for clarification before we can proceed. First, we might ask what constitutes the best life. Rather than favoring a particular conception of a good life, Savulescu argues that certain traits will promote one's ability to achieve wellbeing, and others will detract from this ability, across various reasonable conceptions of a good life, including desire satisfaction, objective list, and hedonistic. Chronic pain, for example, would undeniably make life worse, while an increase in memory would enhance nearly any plan of life; the ability to control violent outbursts would enable individuals to maintain meaningful social relationships and retain their dignity and independence; intelligence would promote wellbeing whether it is through being able to imagine various pleasures, or choosing the most effective means to satisfy one's ends (2001, 420). According to Savulescu, parents are obligated to select children whose traits can be expected to best achieve wellbeing according to such theories, where selection is possible.

Savulescu further argues that enhancements that close off reasonable goods and opportunities are *not* acceptable. It is important to parse this carefully because at first glance, it seems obvious that parents will prize and pursue certain goods at the expense of others. Some parents might emphasize being prudential and moral agents, thereby encouraging development of virtues necessary for coping with the contingencies of life and treating others with respect and kindness. Such parents might enlist their children in

3. An earlier version of this principle appears in Savulescu (2001).

religious training or participation in community service. Other parents might be invested in developing intellectual capacity and have their children participate in chess clubs or math teams.⁴ How then should we understand this restriction? Proponents of genetic enhancement appeal to Joel Feinberg's (1980) 'rights in trust', which ensure "the child's right to an open future". These rights must be preserved for children even though they lack the capacities to exercise them now.⁵ The idea here is that children have a negative claim against parents that they not intervene in their lives in a way that forecloses on their opportunities in the future. Thus, even though a child cannot now exercise her right of reproduction, parents cannot take measures to deprive her of this possibility. Such prohibitive conditions are echoed in Buchanan et al.:

it would be wrong for parents substantially to close off most opportunities that would otherwise be available to their children in order to impose their own particular conception of a good life or in order to continue their own community that is committed to that conception of a good life.⁶

Larry Herzberg (2007) notes that the obligation to enhance should be constrained by an autonomy-respecting condition that also

prohibits any enhancement that would result in the creation of a person with fewer rationally desirable life-options, or with less of a cognitive ability to choose between them, than would otherwise have been the case. For instance, it prohibits enhancements that would make particular occupations harder for the person with the enhanced trait to choose or pursue, even if it would make other occupations easier. (100)

While parents may encourage certain activities or life styles, they cannot intervene in ways that eliminate the ability to exercise certain rights. For example, rights in trust

4. Such differences are highlighted by Buchanan et al in their discussion of the extent of freedom parents have in raising their children. (2000, 158).

5. Feinberg's original division of rights is made in the legal context, specifically, to determine whether or not the state should prevent parents from foreclosing children's futures in some way, say, whether they can, due to religious conviction or cultural tradition, prevent their children from attending school or receiving blood transfusions. Feinberg's distinction has, however, also proved useful and relevant to the ethical boundaries between parent involvement and a child's autonomy, and has been utilized by a number of authors in the area including but not limited to Dena Davis, Michael Sandel, William Buchanan and Norman Daniels.

6. Buchanan et al. *From Chance to choice*, 170

would prohibit a deaf couple from using genetic therapy or selection to produce a deaf child. While this would create a strong bond within family members and even the deaf community, the act will nevertheless foreclose certain opportunities for the child. In her (2010), Davis explains that the child “will have only very limited options to move outside of that culture.... [confining] her forever to a narrow group of people and a limited choice of careers ” (82). More needs to be said about the type of rights involved and what, specifically, counts as infringement, but the principle’s prohibitive message is sufficiently clear for our purpose.

A second point of clarification concerns Savulescu’s use of ‘significant moral reason’, which is to be read as a qualified obligation. More precisely, “when the obligation to have such a child is not overridden by sufficiently strong opposing moral reasons, it will be true that parents ought, all things considered, to select the most advantaged child” (2009, 278). We can acquire a more precise idea of what counts as competing normative reasons, from Savulescu’s (2009) criteria for an *ethical* enhancement. An ethical enhancement is not only in the child’s best interest in that it is expected to be reasonably safe and beneficial, but also cannot unreasonably restrict the range of possible lives open to her, directly harm others through excessive costs, or reinforce and increase unjust inequality and discrimination.⁷ These clarifications allow us to define Savulescu’s position more perspicuously: Parents have an all-things-considered obligation to use ethical enhancements.

III. The Objection from Harm and Savulescu’s Response

As shown above, proponents of pre-natal genetic enhancements cautiously limit their approval to ethical enhancements: those that can be expected to benefit the child according to different conceptions of a good life without causing harm to the child, family or wider society. Such constraints address a number of concerns that have previously been raised against the use of genetic technology. A pluralistic understanding of a good life, along with deliberate attention to the child’s future interests respond to concerns about the return of 20th century eugenics and threats to the child’s autonomy. However, these conditions do not speak to a different sort of worry. Despite careful efforts to secure the child’s wellbeing, enhancements may misfire by seriously harming the child. This is especially troubling for non-disease cases, because such children might have lived entirely normal, albeit unenhanced, lives. Suppose scientists discover a correlation between

7. For some discussion of possible exceptions, see Savulescu 2010.

enzyme E and greater attention span, which by and large, has been identified as a desirable cognitive skill. Geneticists have isolated the gene responsible for the production of this enzyme and have discovered a way to introduce the gene to existing embryos. *P* opts to enhance her child in this way, reasoning that this will increase the likelihood of the child's wellbeing without creating significant hardship for others. The enhancement is ethical in Savulescu's terms. Suppose that in so doing, *P* inadvertently removes certain other desirable traits such as spontaneity, or worse, generates a predisposition to cancer. *P* has harmed her child by making her worse off than she otherwise would have been, despite meeting the conditions of an ethical enhancement. Note that this is not akin to cases in which parents have intervened to treat or prevent some catastrophic illness. If *P* had done nothing at all, the child would have a species-normal attention span and been spared a potentially deadly illness. Contra Savulescu and others, it seems that *P* ought not to have used the ethical enhancement.

Savulescu (2001, 2006) defends genetic intervention by distinguishing selection from enhancement. Enhancement involves the alteration of particular genes, through gene therapy (gene insertion into gametes or embryos) or gene surgery, (where undesirable genes are deactivated) (Buchanan et al. 2000). Selection, on the other hand, is carried out through analysis of the embryo via in vitro fertilization or pre-natal testing through chorionic villus sampling, amniocentesis and ultrasound. The difference is that enhancement operates on a given embryo, whereas selection is a screening process that enables parents to select the embryo that contains traits they desire. Suppose that Lisa was selected via in vitro fertilization because she possessed genes correlated with increased attention span. But Lisa also has a genetic predisposition to cancer. According to Savulescu, the parents have not harmed Lisa. If instead the parents had chosen an embryo with a different genetic blueprint (one who both lacked the desirable traits and cancer), then *Lisa* would not have been born. So long as Lisa's life is worth living at all, the parents have not harmed Lisa because they have not made Lisa's life worse than it would have been.⁸ Suppose instead that parents altered Lisa's genetic structure in a way that enabled greater attention span, but in so doing caused Lisa to develop cancer. Here, Savulescu argues that parents harmed Lisa by making her life worse off than it would

8. McMahan (1997) further points out that because Lisa's life is worth living, parents have benefitted the child: "Indeed, if it is bad for a child to be caused to exist with a life that is not worth living, then it seems that by parity of reasoning, it should be good for a child to be caused to exist with a life that is worth living". (12)

have been without the intervention. Since harm-to-the-child is what is at issue, selection appears immune to this objection whereas enhancement does not.

Savulescu adopts what he (following Parfit) calls a *person-affecting* view of harm, according to which a person is harmed by an act if she is made worse off than she would have otherwise been. The person-affecting view of harm is plausible because we tend to think that harm occurs when (i) a particular individual or set of individuals is affected and (ii) when this individual or set of individuals is comparatively worse because of the act.⁹ Parfit (1973, 1984) and subsequent writers have presented cases in which it seems right to say that harm occurs even if a particular individual or set of individuals is not made worse off, generating the *non-identity problem*. In his (1984), Parfit discusses a 14-year old girl who is told that if she waits a few years to have her child, she will give it a better life (358). She decides to ignore this advice and have the child, who has a difficult start, one that is much more difficult than if she had waited a few years. While it seems right for the girl to have waited, we cannot say that her act was worse for *her* child. This is because had she waited, this child would not have existed at all, and because this child's life is worth living, the girl's decision was not worse for this child. The person-affecting view of harm results in the odd conclusion that the mother has not committed harm.¹⁰

The worry is also illustrated in McMahan's (1997) 'first preconception variant':

A couple are considering having a child but suspect that one of them may be the carrier of a genetic defect that causes moderately severe mental retardation or cognitive disability. They therefore seek to be screened for the defect. The physician who performs the screening is negligent, however, and assures the couple that there is no risk when in

-
9. Savulescu offers a counter-factual explanation of harm: *A* has been harmed when *A* is made worse off than she would otherwise have been if the act had not been performed. This characterization seems problematic. Suppose that Max stabs Lisa with a knife. I'm inclined to think that Max harmed Lisa even if, given Lisa's circumstances, if Max had not stabbed her, she would have been stabbed by someone else, or suffered a knife accident. While I think that this is a relevant concern, I will not pursue it here because the focus of my argument is not this aspect, but what is identified in (i): the target of harm.
10. Curiously, Savulescu argues that his own remarks on selection "echoes" and, draws from, Parfit's non-identity problem. I take it that this is because selecting an embryo with better prospects for wellbeing is similar to waiting to have a child with better prospects for wellbeing. However, Savulescu's own view crucially departs from Parfit's in that when an act results in a worse outcome that involves a different person, Savulescu's view cannot acknowledge that this act has produced harm, or at least does not adequately appreciate the severity of the harm.

fact the man is a carrier of the defect. As a result, the couple conceives a child with moderately severe cognitive impairments.

If the physician had screened the couple properly, a single sperm would have been isolated and genetically altered to correct the defect, which would have been combined in vitro with an egg drawn from the womb. This would have resulted in a child without cognitive disability, and, importantly, have been an altogether different child. So long as the child naturally conceived has a life worth living, the negligent physician's act was not worse for *this* child, and therefore, according to the person-affecting principle, the physician has not committed harm.

These cases suggest that a person-affecting conception of harm is insufficient. We need a different conception to accommodate our intuition that the physician and 14-yr-old caused harm. Parfit proposes Q:

If in either of two possible outcomes the same number of people would ever live, it would be worse if those who live are worse off, or have a lower quality of life, than those who would have lived. (360)

McMahan calls this the *Impersonal Comparative Principle*, comparative not because it makes a single individual worse off, but rather because a different course of action would have secured the same goods (same number of lives) without the suffering that resulted from the original act:

The objection to causing the [cognitively impaired] child to exist is that it was possible to cause a different child to exist whose life would have contained at least as much good but less of what is bad—in particular, less overall suffering (1997, 32)

This principle also accounts for the harm committed by the 14-yr-old, as she could have had a child without the suffering experienced by the original child. (It is important to highlight that the view is restricted to same number choices, and therefore does not claim that it is wrong or worse to bring about a child with disabilities when it was not possible for a couple to have a child without disabilities.¹¹) Impersonal harm appears plausible, and subsequently renders selection equally susceptible to the objection from harm.

11. While McMahan endorses the impersonal comparative principle, which he attributes to Parfit, he does not support Parfit's *no difference view* between person-affecting and impersonal harms. Instead, he argues that impersonal harms can only be as strong as, but never greater than, person-affecting harms.

Savulescu (2006) contends that personal harms will always be worse than impersonal harms. Imagine for instance, a child who discovers that her genes have been altered. She might reason that were it not for her parents' intrusion, she would have existed quite differently, thereby resenting her parents and developing an embittered attitude toward life. This stinging realization, however, cannot occur to an individual that has been selected. Where does this leave us? Savulescu argues that selection is *more* defensible against the harm objection, arguing in fact that selection is more immune "regardless of how misguided the parents' genetic choices may turn out to be, provided only that the child has a life worth living" (164).

Note that Savulescu has stopped short of fully addressing the problem of harm, instead arguing that enhancement is more problematic than selection. This is unsatisfying for two reasons. First, the difference in suffering Savulescu cites will be small, or at any rate insignificant, in the type of case presented above, where genetic intervention resulted in serious harm. Suppose that Lisa₁ was a product of selection and Lisa₂ a product of genetic therapy, and both develop cancer. It seems that Lisa₂'s resentment towards her parents will pale in comparison to the physical and emotion toil of fighting cancer, making both their suffering roughly equal. Further, Lisa₂ may not even be aware that her parents' intervention was responsible for her disease, which would also make their suffering comparable. Second and more importantly, Savulescu has failed to explain why parents are obligated to perform ethical enhancements despite causing such suffering.

IV. A Subjectivist Solution

We cannot deny that genetic intervention caused significant harm in both cases. The question now is whether such harm removes a parent's obligation to ethically enhance. The problem is best viewed as an application of a much broader debate in metaethics concerning the nature of obligation. According to objectivists, one's obligation, what one ought to do, is determined by all relevant normative facts. Subjectivists, on the other hand, maintain that one's obligation is determined by normative facts that are in some way epistemically accessible to the agent. Proponents of genetic enhancement support their view on the basis of expected benefits, adopting a thoroughly subjectivist perspective. Consider Savulescu and Kahane's (2009) remarks on decision-making:

When we make decisions, the option we should choose is the one which maximizes expected value. In the case of selection and reproductive decision-making, the outcome of interest should be how well a new person's whole life goes, that is, well-being. PB thus states

that we have reason to select the child who is expected to have the most advantaged life. We cannot know which child will have the best life. (278)

John Harris (2010) writes, “parents would act ethically if they were to attempt to achieve such an [enhancement] for their children”. Here is Buchanan et al. (2000): “It is morally desirable or morally good for parents to use a variety of means, including genetic interventions to attempt to produce the best children possible” (162).¹² Even given our best efforts, children may in fact *not* live a better life as a result of such interventions, as Savulescu and Kahane note:

Those born with the greatest gifts and talents may squander them while those born to great hardship may overcome enormous obstacles to lead to the best of lives. It is not surprising that there are such limits on what prospective parents can reasonably hope to achieve through genetic selection (278)

In contrast, the objection from harm is rooted in the view that one’s obligation is determined by *all* relevant normative facts, including the actual outcome, reflecting objectivism. Because in the cases under consideration the resulting individuals were made worse off, opponents argue that parents are obligated to not intervene.

Whether such harm removes an obligation to ethically enhance, then, will depend on the conception of obligation we favor.¹³ If we reject objectivism, then such harms—facts that could not have been reasonably predicted—cannot count against the obligation to enhance. I want to suggest that we have good reason to adopt subjectivism in the context of parental obligation, if not more widely. It is in this way that obligation can

12. Buchanan et al. do not endorse an *obligation* to intervene, but accept that the permissibility of intervention is based on our attempts to produce the best children possible, not that such efforts *do* result in the best children possible.

13. As stated above, this is an application of a meta-ethical debate regarding the nature of moral obligation, specifically (a) whether moral obligations are restricted to reasons an agent possesses and if so, (b) what kind of epistemic constraints are relevant to possession. Much has been said historically and in more recent literature: Graham 2010, Lord forthcoming, 2010 and 2013, Schroeder 2008, 2009. My concern is strictly in its application to genetic enhancements and therefore will have limited bearing on the broader debate. However, such concrete application is helpful in both a theoretical and practical sense. This particular context testifies to the strength of the subjectivist position, and therefore goes some way towards furthering the dialectic. Second, some of the points raised in the larger meta-ethical discussion seem particularly well suited to the situation parents face in making reproductive decisions.

serve as a practical guide in deliberations of what we ought to do. When we have strong evidence that placing children in highly reputed schools will enable them to achieve their goals, we see this as reason for doing so. Because we justifiably believe that vaccinations will protect them from disease, we immunize. Buchanan et al. (2000) point out, that “parents are expected to... keep their children away from drugs, from street crime, from hazardous play.... heed nutritional and dietary concerns” because we have “some claims to know scientifically what is best for children” (157). Parental duties are, in other words, constrained by facts that can shape and direct our moral deliberation, those which are epistemically accessible. The importance of epistemic access to obligation is nicely captured in Errol Lord’s recent defense of subjectivism.

Lord argues that an obligation to ϕ requires possessing the right reasons in favor of ϕ -ing. His argument is based on two premises:

1. S has an obligation to ϕ only if S has the ability to ϕ for the right reasons.
2. S has the ability to ϕ for the right reasons only if S possesses the right reasons. (9)

He argues, first, that an agent has an obligation to ϕ only if she can ϕ for the right reasons, what he calls the *right reasons ability condition*. A right reason is understood in terms of normative facts, or facts that recommend actions. What does it mean to act for the right reasons and why is this ability necessary for obligation? Acting for the right reason requires that the agent’s act be explained and justified by her reason and that she is appropriately sensitive to the relationship between the fact and the act. Suppose that Lou believes that a nearby private school will better prepare her child for college than the public school alternative, and therefore decides to send her daughter there. But Joan decides to send her daughter there because it affords her a certain prestige. Note that the child’s wellbeing both explains and justifies Lou’s act. Joan’s act is explained but not justified by her reason. Further, if it were not the case that private school better prepared her child, Lou would probably not enroll her daughter there, though Joan would. For these reasons, only Lou has acted according to the right reasons. Why is this ability a necessary condition for an obligation? It must be possible for the agent to act according to her obligation in a non-lucky or non-accidental way. If we reject the right reasons ability condition, and instead accept that an agent’s obligation to ϕ does *not* require an ability to ϕ for the right reasons,

there would be cases where you ought to ϕ even though the reasons that make this true cannot get any legitimate grip on you—i.e., they cannot move you in a non-accidental way. In these cases you will *have* to get lucky in order to do what you ought. (11)

If one has an obligation to ϕ , but it is *not* the case that one can ϕ for the right reasons, then the rationale for ϕ -ing is completely out of one's epistemic grasp. This means that if she were to ϕ , she would do what she ought as a matter of pure chance or luck. This is doubly suspect because in such cases it is *impossible* to praise or assign credit to the agent despite doing exactly what she ought to do.

Lord goes on to argue that acting according to the right reasons requires *possessing* the right reasons, characterizing possession as bearing a positive epistemic relation to the reason. S possesses a reason *r* if she is in a position to justifiably access *r* without a significant change in her epistemic situation. What this amounts to in the case of inferential beliefs is that you could come to have a justified belief that *r* if you could and did attend to the contents of your existing beliefs and formed the belief in the right way. In the case of non-inferential beliefs, you have some experiences such that if you could and did attend to certain features of those experiences, and formed the belief *r* in the right way, you would have a justified belief that *r*. Note that this means that the agent does not in fact have to know or believe that *r*, but only that she would be able to come to believe it.¹⁴ Possession is a necessary condition for acting for the right reasons since in order to act for the right reasons, the facts must be epistemically available to you. A reason cannot explain and justify your act if you are unaware and cannot become aware of the reason. Lord's argument provides strong support for the view that an agent's obligation is determined by facts accessible to her by emphasizing the intuitive link between obligation and deliberative capacity. That is, an agent must not only be physically, but also deliberatively capable of acting according to her obligation. What, then, does this tell us about genetic enhancements? Insofar as the harm caused by genetic intervention was not foreseeable, it cannot create an obligation to refrain from

14. There are two ways in which my remarks depart from Lord's position. First, Lord presents both an epistemic and what he calls *treatment* condition as requirements for possession, but my focus will be limited to the former. There are two reasons for this. I want to focus on the importance of *epistemic* limitations when it comes to the obligation to enhance. In addition, I don't think that Lord makes a particularly strong case for the need for a treatment condition. Second, Lord argues that agents must be in a position to *know* the relevant reason. I think, however, that this is too strong. I argue only that parents must be in a position to have a justified belief that the enhancement in question is expected to promote the child's wellbeing.

intervention. This is because an obligation to ϕ requires possessing normative facts in favor of ϕ -ing. Harm to the child in the cases considered here, however, was not within the agents' epistemic ken, and therefore cannot affect their obligation to enhance.

While subjectivism appears to correctly incorporate our epistemic limitations, some might worry that it fails to recognize genuine obligations, say in the case of parents who are very poorly epistemically situated:

Influenza

While the flu is typically unpleasant and inconvenient, some strains can be fatal, especially to very young children and the elderly. Researchers uncover a particular gene that helps to immunize the body against all strains of the virus, making it the case that those who have this gene, either naturally or through genetic therapy, will, in all likelihood never have the flu. This not only eliminates some very uncomfortable experiences, but could potentially save lives. Prospective parents *U*, despite having the resources to use genetic enhancements, are not aware of these expected results and cannot become aware, as the news has not reached their rather small and isolated community.

We intuit that *U* has significant moral reason to enhance their child so that she is not susceptible to a potentially lethal virus. Because subjectivism determines obligation according to reasons that are epistemically accessible to subjects, it cannot recognize *U*'s obligation to use such an enhancement. What is more worrisome is if subjectivism might lead to the following sort of case:

Education

Suppose that Parents *A* live in a modern and highly-developed society. But they belong to a smaller community within this society in which schooling is strongly discouraged. There is a deeply entrenched and widespread belief that formal education leads to moral deterioration and inhibits cognitive development by encouraging a sort of intellectual dependency on others. As a result many adults in this community end up without financial security, career options, etc.

We intuit that parents *A*, despite their beliefs, should send their children to school, just as parents *U*, should enhance their child with the anti-flu gene. These cases, however, are importantly different. In Influenza, parents do not have access to information about

the gene. That is, there is nothing about their sensory experience or prior beliefs that can allow them to infer the obligation to enhance their child. Parents in the second case, however, have evidence that depriving children of education decreases the likelihood of achieving a good life. Because they are in a position to form a justified belief about the benefits of schooling, they are morally obligated to provide education for their children according to subjectivism.

One might, however, argue that regardless of *U*'s epistemic position, they are required to affix their child with the anti-flu gene. While it is clear that it would be best to genetically intervene, what parents ought to do will diverge from what is best in these cases. This is because obligation must be constrained by an agent's capacity, which includes both physiological and deliberative abilities. While the endangered status of humpback whales gives us reason to protest whale hunting, which may result in a state of affairs that is overall better than what would occur if we did not protest, such reasons cannot *obligate* us to participate unless we are both physiologically and deliberatively capable of acting on this reason. If one is required, regardless of such abilities, to always achieve the best possible circumstances, this would in fact engender a highly implausible set of 'oughts'.

Another reason that one might believe that subjectivism does not generate the right verdict is because how we might advise parents. It is clear that if *U* were to ask *our* advice, we would counsel them in favor of the relevant enhancement, without first considering their epistemic status. This is nicely brought into relief by Graham (2010):

The question I want answered when I ask myself what my moral obligations are is the same as that which I want answered when, in seeking your help, I ask you what they are; but, to adequately answer me you don't need to consider my evidence concerning my situation; therefore, my moral obligations don't depend on my evidence concerning my situation. (91)

And again here by Thomson (1986):

On those rare occasions on which someone conceives the idea of asking for my advice on a moral matter, I do not take my field work to be limited to a study of what he believes is the case: I take it to be incumbent on me to find out what is the case. (179)

Given that we would *advise* parents *U* to provide the anti-flu enhancements, it might be argued that moral obligation is *not* limited by epistemic constraints.

This conclusion does not follow. Following Lord, I want to say that this has to do with the semantic referent of 'ought', which is not univocal. Rather, 'ought' statements are relativized to different bodies of information. We can ask what parents A ought to do from the body of information salient and accessible to them, and we can ask what parents ought to do from the body of information salient and accessible to us as their advisers. Though different, both will be true relative to our unique information. Our judgments do not establish U's *moral obligation* from their limited epistemic position. This would of course change once U were able to access the relevant truths.

Conclusion

I have argued here that we cannot respond to the problem of harm by denying that certain forms of genetic intervention cause harm or claiming that some cause less harm than others. A more effective strategy is to recognize that this is fundamentally a disagreement about how to approach parental obligation. The objection from harm raises doubts about the obligation to use ethical enhancements only if we accept that obligation is determined by all relevant facts. I have argued, however, that it is more reasonable to accept a subjectivist view of parental obligation. Subjectivism accommodates the intuitive link between deliberative capacity and obligation, and as such, reflects our judgment about parental obligation outside of the context of genetic enhancement. For this reason, actual and unforeseeable harm cannot remove our obligation to intervene.

References

- Buchanan, Allen, Dan Brock, Norman Daniels and Daniel Wilker. 2000. *From Chance to Choice: Genetics and Justice*. New York: Cambridge University Press.
- Daniels, Norman. 2013. "Can Anyone Really be Talking about Ethically Modifying Human Nature." In *Human Enhancement*, edited by Julian Savulescu and Nick Bostrom. Oxford: Oxford University Press.
- Davis, Dena. 2010. *Genetic Dilemmas: Reproductive Technology, Parental Choices and Children's Futures*. New York: Oxford University Press.
- Feinberg, Joel. 1980. "The Child's Right to an Open Future." In *Whose Child? Children's Rights, Parental Authority and State Power* edited by W. Aidken and H. LaFollette. Totowa, NJ: Rowman and Littlefield.
- Harris, John. 2010. *Enhancing Evolution - the Ethical Case for Making Better People*. Princeton, NJ, USA: Princeton University Press.
- Herzberg, Larry. 2007. "Genetic Enhancement and Parental Obligation". *Philosophy in the Contemporary World* 14 (2): 98-110.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford, Oxford University Press.
- Parfit, Derek. 1976. "Rights, Interests and Possible People." In *Moral Problem in Medicine* edited by Gorovitz, Jemeton, Macklin, O' Connor, Perrin, St. Clair and Sherwin, 369-375. New Jersey: Prentice-Hall Inc.
- Lord, Errol. Forthcoming "Acting for the Right Reasons, Abilities and Obligation." In *Oxford Studies in Metaethics*, vol. 10 edited by Russ Shaffer-Landau.
- Lord, Errol. 2013. "The Importance of Being Rational." Ph.D. diss., Princeton University.
- McMahan, Jeff. 1997. "Wrongful Life: Paradoxes in the Morality of Causing People to Exist". In *Rational Commitment and Social Justice* edited by Jules L. Coleman and Christopher W. Morris, 208-248. New York: Cambridge University Press.
- Sandel, Michael. 2007. *The Case Against Enhancement*. Cambridge: Harvard University Press.
- Savulescu, Julian. 2010. "In Defense of Selection for Nondisease Genes." *The American Journal of Bioethics* 1 (1): 16-19.

- Savulescu, Julian and Guy Kahane. 2009. "The Moral Obligation to Create Children with the Best Chance of the Best Life." *Bioethics* 23 (5): 274-290.
- Savulescu, Julian. 2009b. "Genetic Interventions and the Ethics of Enhancement of Human Beings" In *The Oxford Handbook of Bioethics* edited by Bonnie Steinbock, 516-535. New York: Oxford University Press.
- Savulescu, Julian and Melanie Hemsley, Ainsley Newson and Bennett Foddy. 2006. "Behavior Genetics: Why Eugenic Selection is Preferable to Enhancement." *Journal of Applied Philosophy* 23 (2): 157-171.
- Savulescu, Julian. 2001. "Procreative Beneficence: Why we should Select the Best Children" *Bioethics* 15 (5/6): 413-426.



cognethic.org