PhiMiSci

Philosophy and the Mind Sciences



Nuanced clinical neuroethics

A commentary on Joshua May's Neuroethics: Agency in the age of brain science

Jonathan Pugha (jonathan.pugh@uehiro.ox.ac.uk)

Abstract

In this commentary on Joshua May's Neuroethics: Agency in the Age of Brain Science, I consider some of the implications of May's analysis for clinical neuroethics. In particular, in view of May's appeal to the power of valid consent to deal with some of the issues raised by neuro-interventions, I begin by highlighting that clinical neuroethicists often have to navigate a number of complexities in seeking to facilitate the valid consent of individuals who are potentially subject to various forms of vulnerability. I go on to consider whether May's claim that neuro-interventions can elicit transformative experiences raises any problems for his appeal to consent in this context. In the latter half of the commentary, I consider May's analysis of mental disorder and raise some potential areas of contrast between the concepts of responsibility in criminal justice and autonomy in the medical context. I conclude by suggesting that 'nuanced neuroethics' should not only be vigilant about the sources of empirical evidence it relies upon, it should also attend to the nuances of the particular contexts in which neuroethical arguments are made, and neuroethical concepts deployed.

Keywords

Clinical ethics \cdot Decision-making capacity \cdot Deep brain stimulation \cdot Informed consent \cdot Neuroethics

This article is part of a symposium on Joshua May's book "Neuroethics: Agency in the Age of Brain Science" (OUP, 2023), edited by Carolyn Dicey Jennings.

We are living through an exciting time for the field of Neuroethics. Our understanding of neuroscience, and our ability to use neuro-interventions to modulate the neural processes underpinning key physical, motivational, and affective states continues to develop apace. The growing influence and power of AI promises to lead to further substantial developments in these areas. Moreover, there is an increasing amount of interdisciplinary scholarship of the sort that is necessary to bridge the fields of neuroscience and philosophy in addressing fundamental questions about the nature of moral decision-making.

^a University of Oxford.

In this context, Joshua May's timely book *Neuroethics: Agency in the Age of Brain Science* is a particularly impressive achievement. It succeeds in providing an accessible introduction to the field for newcomers that is comprehensive, and both philosophically and scientifically informed. Importantly though, it also neatly weaves in substantial novel contributions to important debates in the literature that should be taken seriously by those who already have some prior experience of thinking about these issues. Indeed, the five guiding principles that May outlines at the end of the book serve as a memorable statement and reminder of the scholarly values that ought to shape the discipline as it navigates this period of rapid technological and scientific change.

In that spirit, in this commentary I shall offer some friendly critical engagement with some of the original contributions that May makes with respect to my own particular sub-topic of interest, namely clinical neuroethics. Whilst I agree with much of May's analysis relevant to this sub-topic (in chapters 3 and 4), in this commentary I want to suggest that there are two areas in which May's laudable approach to 'Nuanced Neuroethics' may not quite succeed in fully capturing some important complexities in clinical neuroethics. The first pertains to his proposed solution to obstacles to valid consent to neurointervention in chapter 3. The second pertains to his suggested extension of Pickard's concept of 'responsibility without blame' to the medical context in chapter 4.

1 Neuro-interventions, transformative experience, and consent

As May rightly highlights in the precis of the book, neuroethicists have often voiced concerns that certain neuro-interventions could have profound implications for personality, autonomy and agency. Such analyses have *some* empirical grounding; interview studies with recipients of Deep Brain Stimulation (DBS) have revealed that some recipients of this intervention can, unfortunately, experience feelings of self-estrangement following treatment (Agid et al., 2006; De Haan et al., 2017; Schuepbach et al., 2013). Yet, there has been a fair amount of discussion about whether the concerns that neuroethicists have raised in this regard have been overblown, with one widely discussed 2018 paper alleging that the literature on the topic was a bubble that warranted 'deflating' (Gilbert et al., 2021).

One of the main concerns raised in that paper was that the neuroethical discussion of the phenomenon largely invoked poor quality evidence, and thus relied upon "... unsubstantiated speculative assumption in lieu of robust evidence" (Gilbert et al., 2021). May himself does not comment extensively on this particular basis for claiming that concerns about personality change following DBS have been overhyped. However, in light of the principles he defends at the end of the book, I suspect that he would be sympathetic to the claim that 'nuanced neuroethics' should require that we approach the *qualitative* evidence on this topic

with just as much vigilance as we should adopt with respect to neuroscientific evidence in Neuroethics more generally. That said, colleagues and I have elsewhere defended the view that neuroethicists who are sufficiently aware of the limitations of the existing qualitative evidence pertaining to personality changes can still legitimately engage in philosophically and empirically-based mechanistic reasoning about particular case studies of this sort of phenomenon (Pugh et al., 2021).

May's own basis for claiming that neuroethical concerns about personality change following neuro-interventions are overblown is that there are substantial parallels between the changes we might observe following the use of neuro-interventions (like DBS), and other widespread transformative experiences that people undergo in everyday life (such as puberty, becoming a parent, and bereavement). In view of this, he suggests that neurointerventions should not be understood to pose a special ethical problem; rather, the actual problem that they raise concerns medical hubris, and the implications that this has for appropriate risk-benefit analyses. The solution here, May suggests, is to ensure that health professionals are honest about the limitations of our general understanding of neurointerventions, and to allow individuals to make their own autonomous choices about whether to undergo them (May, 2023, pp. 87–88).

I am sympathetic to many of the contours of May's argument here; however, I want to suggest that his solution elides two complexities that warrant further consideration for readers interested in this particular topic in the book. Whilst May is quite right to highlight the challenges that medical hubris poses for the prospect of patients making autonomous treatment decisions, there are a number of other further challenges that cannot be solved by simply enjoining physicians to be honest about what is known about an intervention's risks and benefits. More specifically, I shall suggest that there are at least two other important challenges to consider when obtaining valid consent in this context, besides ensuring adequate risk disclosure. These include (i) the particular vulnerabilities of many candidates for the clinical application of neurointerventions and (ii) the potentially transformative nature of these intervention, something that May himself highlights in his discussion. I shall discuss each in turn.

To his credit, May acknowledges in the book that not all patients are able to make completely free treatment choices with respect to neurointerventions. He illustrates the point with the 1874 example of Robert Barthalow's treatment of Mary Rafferty. Rafferty was a "patient with an open wound in her skull, apparently dying from brain cancer", who Batholomew convinced to undergo crude experimental forms of neurostimulation. May claims in his discussion that "Rafferty was not only subjected to gratuitous suffering, she also did not provide valid consent" (May, 2023, p. 69). The example is a striking one, but it bears re-emphasis that there are often very subtle and complex questions to consider about potential threats to valid consent that contemporary clinical neuroethicists can encounter. For example, with respect to Deep Brain Stimulation, we are often considering an intervention which is typically either being considered as a 'last resort' treatment in clinical

contexts for patients who have few other options (Stevens & Gilbert, 2021), or as a highly experimental intervention in a research context for individuals living with chronic treatment-refractory conditions. Moreover, we might occasionally be considering the use of the intervention for individuals living with complex psychiatric conditions which *may in some cases* have implications for their decision-making capacity. I shall return to the implications of May's discussion of this latter point in chapter 4 below.

To be clear, the point I am raising here is *not* to suggest that the considerations adverted to in the preceding paragraph preclude the possibility of individuals from providing valid consent to neurointerventions. The point is rather that clinical neuroethicists often have to navigate a number of complexities in seeking to facilitate the valid consent of individuals who are potentially subject to various forms of vulnerability (Bell et al., 2014). True, facilitating consent necessitates honest communication about risks and benefits; but it may also require (among other things) (i) consideration of the effects that desperation can have on decision-making (Dunn et al., 2011; Swift, 2011), (ii) measures to avoid the therapeutic misconception in research contexts (Fisher et al., 2012) and (iii) considerations of the various complex implications that symptoms of psychiatric disorders can have for autonomous decision-making. Accordingly, despite the theoretical neatness of solving the ethical problem posed by medical hubris associated with neuro-interventions by appeal to the apparatus of valid consent, we should not make the mistake of assuming that this is a practically straightforward and exhaustive solution.

The second point I want to raise is that the appeal to consent might also not be as straightforward a theoretical solution as May's discussion appears to suggest, when taken in conjunction with his appeal to L. A. Paul's conception of transformative experiences elsewhere in the chapter. I take it that part of May's strategy in suggesting that changes following neurointerventions can constitute a transformative experience is to support the idea that such changes are not particular special; after all, transformative experiences are a common part of life. However, the fact that transformative experiences *per se* are widespread may not alone establish that the type of transformative experience associated with neurointerventions does not raise special problems.

Indeed, there is a lively debate about whether the phenomenon of transformative experiences is compatible with conventional models of informed consent (Egerton & Capitelli-McMahon, 2023; Villiger, 2024a). Briefly, one challenge in this regard is to explain how an individual could make a sufficiently informed decision to undergo a procedure that will elicit a transformative experience, if the nature of such an experience is (by definition) epistemically inaccessible prior to undergoing it. Another question to consider is whether the values grounding a token of valid consent should carry moral weight in a decision to undergo an experience that will serve to radically change those very values. Pertinently for the context of neuroethics, theoretical debates about these challenges are currently being played out in discussions about the provision of psychedelic- assisted therapy. Here, whilst

some authors suggest that the transformative nature of psychedelics is compatible with the possibility of valid consent (Villiger, 2024b), others suggest that they provide us with reasons to employ enhanced consent procedures (Smith & Sisti, 2021) (similar in kind to those that have previously been suggested in the context of DBS, Maslen et al., 2018). More sceptically, Jacobs has claimed that the transformative nature of the psychedelic experience suggests that truly informed consent in the psychedelic context may not be possible (Jacobs, 2023).

Given the apparent dialectic reasons for which May invokes the concept of transformative experience in his discussion, I suspect that he is not particularly concerned about the challenges that they have been understood to raise for our understanding of informed consent. I am inclined to agree that we should not be overly concerned about the implications of transformative experiences for consent in general. Nonetheless, these challenges should not be dismissed merely on the basis that life is full of transformative experiences that do not appear to preclude the possibility of valid consent. One reason for this is that there can be quite different types of transformative experience, and the transformative experiences we undergo in everyday life may not all raise exactly the same kinds of problem for valid consent.

For example, one striking way in which such experiences associated with neurointerventions can differ from others is that they often do not arise gradually or cumulatively in the manner that Carel & Kidd (2020) suggest is typical of more common forms of transformative experience (such as puberty for example). Second, as Jacobs highlights in his sceptical discussion about the possibility of truly informed consent in the psychedelic context, consent transactions here occur in the context of an "... asymmetrical, professionalized relationship between a fiduciary and a vulnerable person, governed by a duty of care" (Jacobs, 2023, p. 8). Accordingly, there is perhaps some reason to suppose that the requirements of consent that are invoked in the context of transformative experiences in everyday life may not be entirely co-extensive with those that are invoked in the context of clinical medicine and research.

These are not knock-down objections to May's analysis and they are not intended to be. Instead, my hope is to highlight areas in which there are further depths to plumb for readers of May's analysis on these points. In the next section of the commentary, I will turn to May's discussion of mental disorder and its implications for clinical neuroethics.

2 Mental disorder, responsibility and decisionmaking capacity

Historically, mental disorders were often understood to necessarily connote irrationality (Bortolotti, 2013), and there are unfortunately numerous examples in which individuals living with mental disorder were denied the authority to make

decisions about their own lives on this basis. May's discussion of mental disorder in chapter four provides an arsenal of powerful arguments for why this 'standard naïve' should be rejected in favour of a more nuanced view. According to May's nuanced view, psychiatric symptoms only sometimes reduce agency, and both 'neurotypical' and neuro-atypical' individuals alike can be understood to lie somewhere on a 'cognitive continuum'.

Again, I have little to add about the broad contours of May's overall argument here. Instead, I want to make some brief comments about its implications for clinical neuroethics, and some potential areas of contrast between the concepts of responsibility in criminal justice and autonomy in the medical context that May alludes to in his discussion.

May largely develops his nuanced position over the course of discussing whether those living with mental disorder are less morally responsible for immoral actions they might perform. Notably though, in the conclusion of chapter four, he also suggests that the nuanced approach he endorses is dominant in both legal and medical settings, pointing out that the capacity of patients to make "an informative and autonomous decision is not determined simply by categorizing them as having a mental disorder" (May, 2023, p. 117). This latter claim is borne out in the Mental Capacity Act of England and Wales; section 2 (3)[b] of that act explicitly stipulates that a lack of decision-making capacity cannot be established merely by reference to a particular condition (Mental Capacity Act 2005, 2005). However, it is perhaps worth noting that, at least in some jurisdictions, these sorts of capacity assessments may not always play the same legal role for patients diagnosed with a mental disorder, as they do for patients without such a diagnosis.

To illustrate, in England and Wales, if a patient is not diagnosed with a mental disorder, then their capacity to make a treatment decision (as outlined in the Mental Capacity Act) is paramount to ascertaining whether over-riding their decision can be legally justified (Szmukler & Weich, 2017). In contrast, under the Mental Health Act (and other examples of conventional mental health legislation), which can be invoked with respect to the treatment of patients diagnosed with a mental disorder, "decision-making capacity plays little or no part in the initiation of detention or involuntary psychiatric treatment" (Szmukler & Kelly, 2016). This continues to be a source of significant controversy, given the crucial role of capacity assessments in establishing the lawfulness of medical treatment in nonpsychiatric domains (Szmukler & Weich, 2017). Although May does not comment on this particular issue, it seems to me that his emphasis on the continuity between neurotypical and atypical agency in this chapter would likely be welcomed by supporters of a 'fusion' approach to law in this area, which would seek to govern "nonconsensual treatment of both 'physical' and 'mental' illnesses, based on incapacity principles" (Dawson & Szmukler, 2006; Szmukler & Kelly, 2016).

One of the implications that May does highlight in his discussion is that his arguments can help to both humanise and empower psychiatric patients, and allow us to recognise an individual living with mental disorder as "a fellow person

who can be held accountable" (May, 2023, p. 114). This is obviously a welcome implication; but I want to note that here (as elsewhere in the book), May implicitly appears to be suggesting that his arguments should be understood to have implications for both our understanding of moral responsibility *and* autonomy. I think this can sometimes be a justifiable approach, but I want to suggest that there is also room for a degree of separation between these concepts, and for caution in moving between them. Although it is quite common in to treat the two concepts co-extensively in philosophical discussions (Fischer, 1999, p. 58), some theorists have been critical of this phenomenon (Oshana, 2002). Indeed, I shall now suggest that May's treatment of blame on his nuanced view provides a practical example of how these two concepts can raise quite different moral questions.

Whilst the nuanced view accommodates the thought that neuro-atypical individuals can often be morally responsible, May takes care to clarify that this conclusion does not entail that they (or indeed morally responsible neurotypical individuals) should be *blamed* for their conduct. Drawing on Pickard's work on responsibility without blame, May suggests that blame can frequently be mitigated by various factors that do not suffice to wholly undermine the individual's moral responsibility. He notes though that:

None of this implies that pathological symptoms or behaviour are desirable or morally acceptable . . . In practice, mitigation of blame just comes so cheaply we should expect to see it as warranted, even when one remains accountable. (May, 2023, p. 116)

Given the suggestion elsewhere that his nuanced view is applicable in both medical and legal contexts, the nuanced view's implications for responsibility and blame raise a number of questions about the application of the view to the medical context. First, should the claims about 'responsibility without blame' be understood to similarly extend to the medical context? Second, if so, what is the appropriate analogue of the concept of 'responsibility without blame' for autonomy and decision-making capacity? Would the thought be that a person can have some degree of autonomy with respect to their treatment decision, but that their decision should not be taken to have complete authority with respect to the matter of how they should be treated? If that is the most plausible way of making an analogous 'blame without responsibility' claim with respect to autonomy, then there is of course scope to query whether that claim is plausible.

Third, assuming it is plausible, there is an important question about the costs associated with diminishing a (somewhat) autonomous individual's decision-making authority.

May does not detail precisely what the moral costs associated with the mitigation of blame for somewhat morally responsible individuals are, but the quote above suggests that he takes them to be minimal in so far as the mitigation of blame 'comes cheaply'. In contrast though, it seems that denying an individual the authority to make their own medical treatment decisions always involves considerable

moral costs. Unwanted medical treatment often leads to considerable distress, and the infringement of the individual's bodily rights, even if it is deemed to be in the patient's best interests. Even if these costs might ultimately be justified by certain countervailing considerations in some cases (for instance, if we believe that the decision is not sufficiently autonomous to warrant ultimate authority), these costs remain weighty, and I suggest, different in kind to those that we countenance in mitigating blame.

Of course, there can also be considerable moral costs on the other side of the balance; that is, there can be costs associated with affording individuals the authority to make potentially harmful treatment decisions, and these costs are particularly salient when the individual lacks the capacity to make that decision in what might be understood to be a sufficiently autonomous manner. My aim here is not to suggest anything about how this balance should be struck - the point is rather that clinical neuroethics often involves the need to make decisions about how best to navigate this tightrope, and that the complexity of the task partly lies in the fact that the stakes are so high. The costs of either kind of error are of considerable moral significance. In this regard, there might be scope for important differences in the implications that May's nuanced view has in clinical neuroethics, and those that it might be understood to have in our practices of blame (on the assumption that the mitigation of blame 'comes cheaply').

To be clear, this is not a criticism of May's arguments per se, or his analysis of mental disorder. Indeed, I applaud the scope of May's discussion of mental disorder, and his wide consideration of the ethical implications of his arguments. Neuroethical scholarship must unavoidably address issues that pervade an increasingly wide range of domains and practical contexts, and May's analysis rises admirably to that challenge. However, I do think there are some considerable challenges in extending the concept of 'responsibility without blame' into the medical context. This illustrates a broader point, which is that nuanced neuroethics should take care to attend to the nuances of the particular contexts in which neuroethical arguments are being made - it may be a mistake to assume that claims about a concept in one neuroethical context (such as blame and responsibility in the law) can be straightforwardly extended to another (such as consent in clinical ethics). More generally, perhaps the most general lesson here is that things in neuroethics can be perhaps even more nuanced than May himself suggests.

References

Agid, Y., Schüpbach, M., Gargiulo, M., Mallet, L., Houeto, J. L., Behar, C., Maltête, D., Mesnage, V., & Welter, M. L. (2006). Neurosurgery in parkinson's disease: The doctor is happy, the patient less so? In P. Riederer, H. Reichmann, M. B. H. Youdim, & M. Gerlach (Eds.), *Parkinson's disease and related disorders* (pp. 409–414). Springer Vienna. https://doi.org/10.1007/978-3-211-45295-0_61

Bell, E., Racine, E., Chiasson, P., Dufourcq-Brana, M., Dunn, L. B., Fins, J. J., Ford, P. J., Glannon, W., Lipsman, N., Macdonald, M. E., Mathews, D. J. H., & Mcandrews, M. P. (2014). Beyond consent in research: Revisiting vulnerability in deep brain stimulation for psychiatric disorders. *Cambridge Quarterly of Healthcare Ethics*, 23(3), 361–368. https://doi.org/10.1017/S0963180113000984

Bortolotti, L. (2013). Rationality and sanity: The role of rationality judgments in understanding psychiatric disorders. In K. W. M. Fulford (Ed.), *Oxford handbook of philosophy and psychiatry* (pp. 480–497). Oxford University Press.

- Carel, H., & Kidd, I. J. (2020). Expanding transformative experience. European Journal of Philosophy, 28(1), 199–213. https://doi.org/10.1111/ejop.12480
- Dawson, J., & Szmukler, G. (2006). Fusion of mental health and incapacity legislation. The British Journal of Psychiatry, 188(6), 504–509. https://doi.org/10.1192/bjp.188.6.504
- De Haan, S., Rietveld, E., Stokhof, M., & Denys, D. (2017). Becoming more oneself? Changes in personality following DBS treatment for psychiatric disorders: Experiences of OCD patients and general considerations. *PLOS ONE*, 12(4), e0175748. https://doi.org/10.1371/journal.pone.0175748
- Dunn, L. B., Holtzheimer, P. E., Hoop, J. G., Mayberg, H. S., & Appelbaum, P. S. (2011). Ethical issues in deep brain stimulation research for treatment-resistant depression: Focus on risk and consent. *AJOB Neuroscience*, 2(1), 29–36. https://doi.org/10.1080/21507740.2010.533638
- $Egerton, K., \& \ Capitelli-McMahon, H. \ (2023). \ Transformative experience and the principle of informed consent in medicine. \\ \textit{Synthese}, 202 (3), 65. \ https://doi.org/10.1007/s11229-023-04258-4$
- Fischer, J. (1999). Recent work on moral responsibility. Ethics, 110(1), 93-139. https://doi.org/10.1086/233206
- Fisher, C. E., Dunn, L. B., Christopher, P. P., Holtzheimer, P. E., Leykin, Y., Mayberg, H. S., Lisanby, S. H., & Appelbaum, P. S. (2012). The ethics of research on deep brain stimulation for depression: Decisional capacity and therapeutic misconception. Annals of the New York Academy of Sciences, 1265, 69–79. https://doi.org/10.1111/j.1749-6632.2012.06
- Gilbert, F., M. Viana, J. N., & Ineichen, C. (2021). Deflating the deep brain stimulation causes personality changes bubble: The authors reply. *Neuroethics*, 14(1), 125–136. https://doi.org/10.1007/s12152-020-09437-5
- Jacobs, E. (2023). Transformative experience and informed consent to psychedelic-assisted psychotherapy. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1108333
- Maslen, H., Cheeran, B., Pugh, J., Pycroft, L., Boccard, S., Prangnell, S., Green, A. L., FitzGerald, J., Savulescu, J., & Aziz, T. (2018). Unexpected complications of novel deep brain stimulation treatments: Ethical issues and clinical recommendations. *Neuromodulation: Technology at the Neural Interface*, 21(2), 135–143. https://doi.org/10.1111/ner.12613
- May, J. (2023). Neuroethics: Agency in the age of brain science. Oxford University Press.
- Oshana, M. A. L. (2002). The misguided marriage of responsibility and autonomy. *The Journal of Ethics*, 6(3), 261–280. https: //doi.org/10.1023/A:1019482607923
- Pugh, J., Pycroft, L., Maslen, H., Aziz, T., & Savulescu, J. (2021). Evidence-based neuroethics, deep brain stimulation and personality - deflating, but not bursting, the bubble. *Neuroethics*, 14(1), 27–38. https://doi.org/10.1007/s12152-018-9392-5
- Schuepbach, W. M. M., Rau, J., Knudsen, K., Volkmann, J., Krack, P., Timmermann, L., Hälbig, T. D., Hesekamp, H., Navarro, S. M., Meier, N., Falk, D., Mehdorn, M., Paschen, S., Maarouf, M., Barbe, M. T., Fink, G. R., Kupsch, A., Gruber, D., Schneider, G.-H., ... EARLYSTIM Study Group. (2013). Neurostimulation for parkinson's disease with early motor complications. *The New England Journal of Medicine*, 368(7), 610–622. https://doi.org/10.1056/NEJMoa1205158
- Smith, W. R., & Sisti, D. (2021). Ethics and ego dissolution: The case of psilocybin. *Journal of Medical Ethics*, 47(12), 807–814. https://doi.org/10.1136/medethics-2020-106070
- Stevens, I., & Gilbert, F. (2021). Ethical examination of deep brain stimulation's "last resort" status. *Journal of Medical Ethics*, 47(12), e68–e68. https://doi.org/10.1136/medethics-2020-106609
- Swift, T. (2011). Desperation may affect autonomy but not informed consent. AJOB Neuroscience, 2(1), 45–46. https://doi.org/10.1080/21507740.2010.537293
- Szmukler, G., & Kelly, B. D. (2016). We should replace conventional mental health law with capacity-based law. *The British Journal of Psychiatry*, 209(6), 449–453. https://doi.org/10.1192/bjp.bp.116.191080
- Szmukler, G., & Weich, S. (2017). Has the mental health act had its day? BMJ, 359, j5248. https://doi.org/10.1136/bmj.j5248 Mental capacity act 2005, (2005). http://www.legislation.gov.uk/ukpga/2005/9
- Villiger, D. (2024a). Transformative experience. Philosophy Compass, 19(6), e13000. https://doi.org/10.1111/phc3.13000
- Villiger, D. (2024b). Giving consent to the ineffable. Neuroethics, 17(1), 11. https://doi.org/10.1007/s12152-024-09545-6

Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.