

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/155902/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Jefferson, Anneli 2024. Are mental disorders brain disorders? -A precis. *Philosophical Psychology* 37 (3) , pp. 552-557. 10.1080/09515089.2023.2166822

Publishers page: <http://dx.doi.org/10.1080/09515089.2023.2166822>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



## ***Are mental disorders brain disorders? – A precis***

Anneli Jefferson

Author manuscript, the Version of Record of this manuscript has been published and is available at *Philosophical Psychology*: <https://doi.org/10.1080/09515089.2023.2166822>

People hold wildly opposing and very strong views on the question whether mental disorders are brain disorders, and the disagreement is primarily a conceptual one, not one about whether there are, say, relevant biochemical differences in people suffering from mental disorders and those who don't. There are such empirical disagreements, too (see for example recent discussions about the role of serotonin in depression (Erritzoe et al., 2019; Moncrieff et al., 2022), but the disagreement runs deeper than that. There is a long-standing belief that mental disorders must be brain disorders, because whatever psychological dysfunction we find must have some basis in the brain. One could say that the claim follows by metaphysical necessity. By contrast, others hold that 'mental disorders are brain disorders' is a misguided and deeply harmful claim.

In this book, I answer the question whether it is plausible to think that some mental disorders are brain disorders. I focus on the conceptual question of what a brain disorder is and provide a philosophical account of what the claim that mental disorders are brain disorders can plausibly amount to; what it would take for a mental disorder to be a brain disorder. I leave working out the empirical details to the psychologists, psychiatrists and neuroscientists. The account I propose holds that mental disorders are brain disorders if they involve brain dysfunction, and that brain dysfunction need not - and indeed often cannot - be identified as such independently of the fact that it realizes psychological dysfunction. This means that mental disorders qualify as brain disorders if we can identify brain dysfunction, but that brain dysfunction can be derivative of psychological dysfunction.

I start my discussion by looking at an understanding of brain disorders that puts them in opposition to mental disorders, a position I call 'the narrow view', whose most famous proponent is still Thomas Szasz (1960). This narrow view looks to paradigmatic brain disorders such as brain cancer or neurosyphilis as a model for what makes something a brain disorder. It concludes that a condition can only be a brain disorder if the following conditions are met: dysfunction in the brain is specifiable independently of the psychological level; it causally precedes mental symptoms and treatment of the disorder targets the brain directly, for example through medication or surgery. (This specification that the brain is targeted directly is important because psychological treatment through talking therapy targets the brain as well, but does so indirectly.) On this understanding, some conditions with psychological symptoms are brain disorders, but most classic mental disorders, like schizophrenia, depression or addiction are not.

The narrow view is unsatisfactory because both the etiology of brain disorders and the treatment involve a plurality of factors (Kendler, 2012). Furthermore, somatic dysfunction, too, is not specifiable independently of the adverse effects a structural or functional anomaly in the body has for an organism. It's unclear why this should be a suitable standard for brain dysfunction. Having rejected the narrow view, I turn to the claim that all mental disorders are brain disorders, simply because any kind of dysfunctional mental process will have to be realised by some corresponding brain process. This conclusion is too quick, I argue. It presupposes that dysfunctionality at the mental level automatically implies dysfunctionality at the level of the brain. Pointing to the hardware/software distinction and the difference between hardware problems and software

problems, many authors have argued that the same difference holds in the case of the mind/brain (Papineau, 1994). Standards for what counts as dysfunctional may differ between the psychological level and the physiological one.

However, there is an important difference between the relationship between mind and brain and that of software and hardware. Functional specification is given by the creators in the case of artefacts, and there is no readily available analogue in the case of the mind/brain (Jefferson, 2020). Harriet Fagerberg (2022) even argues that the mind-brain relationship is fundamentally different: as mental functions and brain functions have evolved together, mental functions will necessarily have corresponding brain functions.

But this conclusion is too quick. Just as is the case with hardware and software, the relationship between mental dysfunction and brain dysfunction may not be tight enough to justify the claim that every psychological dysfunction involves a corresponding brain dysfunction. The relationship between mental processes and underlying brain processes may not be sufficiently systematic; and it may not be possible to identify a type of brain difference (in comparison to normal function) that corresponds to the mental dysfunction we have identified. This could be because of extreme multiple realization, which may make it impossible to find a difference between brain processes underlying functional and dysfunctional mental processes. The metaphysical purist may object that this is a purely epistemological issue. But where there is no systematic type of brain function that can be identified as having broken down, it does not make sense to speak of brain dysfunction.

Instead, we should say that differences in brain structure or function that we find in mental health conditions are dysfunctional precisely when and because they realize mental dysfunction. Some multiple realization of psychological dysfunction in the brain is permissible on my account, but there need to be systematic patterns of realization which differ from normal functioning. Mental disorders as brain disorders are thus different from paradigmatic brain disorders, where pre-existing brain dysfunction *causes* mental symptoms, as in the case of cancer or syphilis. This gives us a more modest notion of brain dysfunction and disorder, which only requires that we find differences in the brain that are sufficient for realizing psychological dysfunction. A consequence of this view is that for many psychiatric disorders, brain dysfunction will only be specifiable as such because it realizes psychological dysfunction. It is also compatible with proposals by the RDoC (Research Domain Criteria) programme for research on mental health conditions, which suggests that mental disorders are brain disorders because they involve disorders of brain circuits (Insel & Cuthbert, 2015). There are brain dysfunctions that don't realize mental dysfunctions, and it currently looks as though there are mental dysfunctions that are not realised by brain dysfunctions. But when a mental disorder is a brain disorder, we are dealing with a case where a brain dysfunction realizes a mental dysfunction.

The second half of the book is dedicated to fleshing out the account by considering objections and implications. In chapter four, I address two important objections. The first is that brain disorder labels are committed to a problematic reductionism, the second that mental disorders are relational (externalist) in a way that a focus on dysfunction in the brain cannot allow for. My proposal is indeed reductionist if the reduction at stake is explanatory reduction, where we identify phenomena at one level of description with phenomena at a different level of description. However, in debates about psychiatry, this form of reduction is often conflated with a different kind of reduction, where reduction amounts to identifying a single biological or brain-based cause. This form of reduction is not implied by my account. Some authors argue that even explanatory reduction is in principle impossible, but I show that this argument is not well-supported.

The second objection concerns the compatibility of my account with externalism. There are many forms of externalism in the philosophy of medicine, what they share is that they take the relationship between the organism and the external world to be essential in defining or even constituting certain diseases. For example, it is definitional for post-traumatic stress disorder that it was caused by a traumatic event. Another externalist feature of mental disorders is that they are often described, but also labelled as pathological, in virtue of certain reactions being inappropriate given the outside environment. So, for example, depression involves negative emotions that are not merely an understandable reaction to a tragic life event. According to externalists, the distinction between depression and grief lies in the fact that in the case of grief, the emotions are an appropriate response to life events. By contrast, depression involves *disproportionate* low mood (Davies, 2016).

I counter the externalist challenge in a two-fold way: first, I point out that somatic conditions, too, are importantly externalist in that they are sometimes categorized as disordered on the basis of how the body reacts to the external environment. For example, Shane Glackin (2017) draws attention to the fact that photophobia is defined by the eye reacting *too* strongly to light. Furthermore, the central claim that mental disorders are indistinguishable from normal reactions if we exclude the relationship to the environment is under-supported. For example, there are more psychological differences between grief and depression than externalists recognize.

Finally, I turn to some of the moral and practical concerns that arise from regarding a condition as a brain disorder. First, I address the question whether brain disorders excuse people from moral responsibility. While brain disorders can undermine responsible agency, this does not follow automatically from the fact that there was something wrong with the brain, even in those cases where we are dealing with classic brain disorders, such as tumours. Rather, we need to look at how an agent's impulse control, emotions and reasoning capacity are affected by the brain disorder. Only if they are sufficiently affected that the person no longer meets criteria for responsible agency *psychologically described*, is that person no longer responsible (Jefferson, 2022). Knowledge of brain dysfunction can inform such psychological assessments, but it cannot replace them.

There is one further issue that people often raise either in defence or in objection to the 'brain disorder' label: the effect of the label *itself* on the way we perceive people who are labelled as suffering from a brain disorder. While some clinicians have argued for the brain disorder label precisely on the basis that it is destigmatizing (Leshner, 1997), others argue that it is both stigmatizing and harmful to treatment (Satel & Lilienfeld, 2017). It's a well-understood fact that labels affect those labelled (Hacking, 2007), and therefore the good or bad effects a label might have need to be taken seriously. Rightly or wrongly, hearing that a person is suffering from a condition that involves dysfunction in the brain may create the impression that they are unable to do anything to recover from or manage their condition: it may also create the impression that the person is not responsible for their actions.

I look at the empirical evidence concerning the stigmatizing effects that brain-based explanations of mental disorders have on the way others perceive people suffering from these conditions. Findings are mixed, but there is some evidence of stigma and the belief that people suffering from brain disorders are severely impaired in their agency. I look at two possible explanations for such effects: the first is that, in effect, people think of paradigmatic brain disorders when they hear a condition is a brain disorder and draw unwarranted inferences from this, for example that it may only be treated through purely medical means. This can be addressed by broadening the examples of brain disorders we draw on, and by a clearer understanding of the notion of brain dysfunction. The second explanation is that people are intuitive dualists, who think of brain and mind as fundamentally

different. Because of this, brain explanations for behaviour or mental phenomena block or replace mental explanations in our thinking. I argue that in as far as intuitive dualism poses a problem for taking mental processes seriously, it does so for all mental processes and should therefore undermine our notions of human agency and responsibility generally (Greene & Cohen, 2011). We should therefore conclude that if we are not worried that physicalism and neuroscience generally undermine human agency, we should not be worried in the case of psychiatry.

My analysis in this book is purely philosophical, albeit informed by the empirical literature. Whether we should speak of mental disorders as brain disorders will depend on how successful explanatory reduction is. But I hope to have put in place the conceptual resources that give us a better understanding of what talk of 'brain dysfunction' and 'brain disorders' is and is not committed to.

## References

- Davies, W. (2016). Externalist Psychiatry. *Analysis*, 76(3), 290-296.
- Erritzoe, D., Godlewska, B. R., Rizzo, G., Searle, G. E., Lewis, Y., Ashok, A., Howes, O., Passchier, J., Gunn, R. N., Nutt, D. J., Cowen, P. J., Knudsen, G. M., & Rabiner, E. A. (2019). P.361 Brain serotonin release reduced among patients with severe depression: a pet study with [11c]cimbi-36 and d-amphetamine challenge. *European Neuropsychopharmacology*, 29, S258. <https://doi.org/https://doi.org/10.1016/j.euroneuro.2019.09.379>
- Fagerberg, H. (2022). Why Mental Disorders are not Like Software Bugs. *Philosophy of Science*, 89(4), 661-682. <https://doi.org/10.1017/psa.2022.7>
- Glackin, S. N. (2017). Individualism and the medical: What about somatic externalism? *Analysis*, 77(2), 287-293. <https://doi.org/10.1093/analys/anx073>
- Greene, J., & Cohen, J. (2011). For the law, neuroscience changes nothing and everything. In J. Illes & B. Sahakian (Eds.), *The Oxford Handbook of Neuroethics* (pp. 655 - 674). Oxford University Press.
- Hacking, I. (2007). Kinds of People: Moving Targets. In *Proceedings of the British Academy, Volume 151, 2006 Lectures* (Vol. 151, pp. 285-318).
- Insel, T. R., & Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science*, 348(6234), 499-500. <https://doi.org/10.1126/science.aab2358>
- Jefferson, A. (2020). What does it take to be a brain disorder? [journal article]. *Synthese*, 197(1), 249-262. <https://doi.org/10.1007/s11229-018-1784-x>
- Jefferson, A. (2022). Brain Pathology and Moral Responsibility. In J. May & M. King (Eds.), *Agency in Mental Disorder: Philosophical Dimensions*. Oxford University Press.
- Kendler, K. S. (2012). The dappled nature of causes of psychiatric illness: replacing the organic-functional/hardware-software dichotomy with empirically based pluralism. *Molecular Psychiatry*, 17(4), 377-388. <http://dx.doi.org/10.1038/mp.2011.182>
- Leshner, A. I. (1997). Addiction is a brain disease, and it matters. *Science*, 278(5335), 45-47.
- Moncrieff, J., Cooper, R. E., Stockmann, T., Amendola, S., Hengartner, M. P., & Horowitz, M. A. (2022). The serotonin theory of depression: a systematic umbrella review of the evidence. *Molecular Psychiatry*. <https://doi.org/10.1038/s41380-022-01661-0>
- Papineau, D. (1994). Mental Disorder, Illness and Biological Disfunction. *Royal Institute of Philosophy Supplements*, 37, 73-82. <https://doi.org/doi:10.1017/S135824610000998X>
- Satel, S., & Lilienfeld, S. (2017). Calling it 'brain disease' makes addiction harder to treat *Boston Globe*. <https://www.bostonglobe.com/ideas/2017/06/22/calling-brain-disease-makes-addiction-harder-treat/ehaJs5ZYIXpPottG89KOGK/story.html>
- Szasz, T. (1960). The Myth of Mental Illness. *American Psychologist*, 15, 113-118.