Generalised Brain Functions Contradict Themselves

1. Introduction

This paper defends the Dysfunction Account of Disorder against Justin Garson's Generalised Selected Effects Theory of brain function.

My argument has two parts. First, I contest Garson's claim that synaptically selected neuropsychiatric disorders are not dysfunctional. This does not follow from his view. What in fact follows from Garson's view is that neurally selected disorders are *both* functional and dysfunctional. I will then argue that this consequence of Garson's theory is sufficient grounds to reject it outright.

2. The Traditional Theory

In his 2019 book 'What biological functions are and why they matter', Garson contrasts his 'Generalised Selected Effects Theory' with the standard evolutionary account, which he terms the 'Traditional Selected Effects Theory' (from here, the 'Generalised Theory' and the 'Traditional Theory').

The Traditional Theory states:

> A *function* (F) of a trait (T) in an organism (O) is an effect of T that increased the inclusive fitness of O's ancestors, in the environment of evolutionary adaptedness, such that T was naturally selected.
>
> (See e.g. Neander, 1991; Millikan, 1989; Godfrey-Smith, 1994)

So, for example, it is a function of my thyroid gland to release appropriate levels of triiodothyronine into my bloodstream because that is the thing which thyroid glands did in the past which led them to be naturally selected.

(From here, I use *ontogenetic function* to refer to effects selected intra-organismically on a developmental timeline, and *evolutionary function* for effects selected intergenerationally on an evolutionary timeline.)

3. The Generalised Theory

The Generalised Theory states:

> A *function* (F) of a trait (T) is an effect that led to T's differential reproduction, or differential retention, in a population.
>
> (Adapted from Garson, 2019, p. 93)

While the Traditional Theory covers only evolutionary functions, Garson's theory manages also to encompass a range of ontogenetic functions by appealing to effects selected by differential retention through ontogeny as well as by differential reproduction through evolution.

The most appealing implication of this move is that it allows Garson to attribute direct proper biological functions to synaptically selected traits in the brain. Synapse selection is a neuroplastic mechanism by which the brain overproduces synaptic connections between neurons and then proceeds to shed the 'useless' ones. According to Garson, the surviving synapses acquire the function of doing whatever it was that caused them to be selectively retained relative to the others – just as the thyroid gland acquired the function of releasing

hormone by increasing the fitness of those carrying its genotype in the evolutionary past (Garson, 2019). Garson thus expands the extent of direct, proper function into the neuroplastic realm.

4. Garson's argument for the Generalised Theory

Garson contends that the same argument which supports the Traditional Theory also supports the Generalised Theory. According to Garson, we were motivated to accept the Traditional Theory in the first place because it made sense of three *prima facie* puzzling features of proper functions: (1) the function/accident distinction, (2) their normativity and (3) their explanatory depth.

(1) The Traditional Theory distinguishes between functions, such as the capacity of the nose to detect odours, and accidental effects, such as its capacity to hold up spectacles. Because the nose evolved to smell (and not to hold spectacles) that is its function. (2) The Traditional Theory also successfully accounts for the 'normativity' of functions. Because a trait's function is determined by its history of selection (as opposed to any current activity) it can fail to perform a function that it is 'supposed' to perform. (3) Finally, we sometimes cite functions as explanations for traits. "Why do Zebras have stripes? For camouflage." The Traditional Theory tells us that a function is an effect which caused a trait it to be naturally selected. Thus, the Traditional Theory can account for functions' apparent explanatory depth (Garson, 2019).

From here, Garson argues as follows:

> The core argument for the [Generalised Theory] is simply a parity of reasoning argument. Consider why we accepted the traditional selected effects view. We did so because it made sense of three big puzzles of function … Since [the Generalised Theory] solves all the same problems, minus an arbitrary distinction, we should accept it.
>
> (Garson, 2019, p. 94)

In other words, the argument which supports the Traditional Theory also, by parity of reasoning, supports the Generalised Theory, and we have no good reason to restrict ourselves to the Traditional Theory's 'arbitrarily' narrow evolutionary scope. We are accordingly forced to accept the Generalised Theory in its place "on pain of inconsistency" (Garson, p. 93, 2019).

What does it mean to accept a 'generalised' theory of function, as Garson contends we must? It is not simply the claim that applying the term 'function' to ontogenetic functions would be correct. Philosophers and biologist already do this in various contexts, so this would be trivial. Rather, Garson is committed to something like the following: evolutionary and ontogenetic selection processes yield functions of *the same kind*. They play – or should play – essentially the same theoretical role. Restricting one's concept of direct proper function to include only evolutionary functions, even in specific domains, is theoretically unprincipled.

Acceptance of the Generalised Theory, Garson then goes on to argue, spells trouble for the Dysfunction Account of Disorder.

5. The Dysfunction Account

The 'Dysfunction Account of Disorder' ('Dysfunction Account', from here) theorises that medical disorder necessitates dysfunction.

It states:

> A *dysfunction* is the inability of some trait T to perform one of its biological functions.
>
> *Medical disorder* necessarily entails dysfunction.
>
> (See e.g. Wakefield, 1992; 2014; Neander, 1983; 1998)

So, for example, hormone excretion is the biological function of my thyroid gland, so its failing to excrete sufficient hormone would constitute a dysfunction, and thus qualify as a medical disorder.

6. Garson's Critique of the Dysfunction Account

In what follows, I lay out Garson's critique of the Dysfunction Account from the Generalised Theory. I focus on the case (offered by Garson) of Substance Use Disorder, however any challenge to the Dysfunction Account based on the Generalised Theory will take a similar form.

A. We should accept the Generalised Theory in place of the Traditional Theory.

B. Following the Generalised Theory, Substance Use Disorder is functional and not dysfunctional.

According to Garson, the Generalised Theory "expands the domain of entities that can that can possess direct proper functions, thereby increasing the likelihood that a given condition is functional, not dysfunctional." (Garson, p. 178, 2019).

On Substance Use Disorder in particular, he writes:

> Though the exact mechanisms [underlying drug addiction] are still a matter of controversy, it's likely that synapse selection is involved. …. Once a synapse is selectively strengthened this way, it comes to acquire a new function. It has the function of causing the behaviour of that led to its differential retention. If that behaviour included seeking out and using drugs, then that becomes its direct proper function.
>
> (p. 180, Garson, 2019; see also Garson & Papineau, 2019)

Precisely what is it that has acquired the proper ontogenetic function of Substance Use Disorder, on Garson's account? There is some ambiguity here, so let us briefly clarify before moving on.

According to Garson, proper functions are proximal: it is the proper function of trait T to yield that effect which is most specific to T or, if you will, that which T can do more or less on its own. So, for example, it is the proper function of the heart to pump; not to circulate blood around the body. According to Garson, "when we identify a trait's function with its

distal effect, we're committing a fallacy of division. The function of circulating blood does not belong to the heart but to a bigger system that includes the heart among its parts" (p. 119, 2019).

It follows from this that the trait T which has the proper function 'Substance Use Disorder' must be the overall neural system which in fact performs this function. Unless some particular synapse is performing Substance Use Disorder all on its own, it would be fallacious to attribute to it this system-level effect. Rather, parts of this system have as their proper ontogenetic functions to contribute (howsoever they in fact contribute) to the overall effect – just as the heart has as its proper function to contribute to blood circulation by pumping.

C. From A and B, we should accept that Substance Use Disorder is functional and not dysfunctional.

D. Substance Use Disorder is a case of mental disorder.

Drug addiction, or Substance Use Disorder as it is termed in the DSM, is a commonly accepted disorder and features in official classifications, so we should *prima facie* accept that it is indeed a case of mental disorder[1].

E. From C and D, we should accept that some disorders are functional and not dysfunctional.

F. The Dysfunction Account states that all disorders necessarily involve dysfunction.

G. From E and F, we should reject the Dysfunction Account.

7. Rebutting Garson's critique

In what follows, I will show that Garson's premise B is false. I will then argue that premise B being false also leads us to reject premise A. In other words, Garson's critique of the Dysfunction Account fails, and this failure exposes an underlying flaw in the Generalised Theory.

7.1. Premise B is false

It follows from Garson's Generalised Theory that Substance Use Disorder is functional. So far, he is correct. However, it does *not* follow from Garson's theory of function that Substance Use Disorders it is *therefore* not dysfunctional. In concluding this, Garson is simply wrong. To see why, we need to know what it is in the brain that in fact performs Substance Use Disorder.

Substance Use Disorder is neurobiologically complex. There are multiple, diverse and to some extent independent structures and circuits implicated in the production of the psychological states and behaviours involved in substance dependence. One neuropsychological theory is that substance abuse is caused and maintained by 'feedback loops' between the reward system, the anti-reward/stress system and the executive systems in the brain (Koob & Simon, 2009). This is supported by functional neuroimaging studies

[1] I grant Garson this for present purposes, but there is debate here (see e.g. Schwartz, 2014; Lemoine, 2013; Matthewson and Griffiths, 2018; Cooper, 2020).

showing associations with neuroplastic changes in several neurological sub-systems of the basal ganglia, the extended amygdala and the orbitofrontal/prefrontal cortex (Koob & Simon, 2009; Koob & Volkow, 2016).

As Garson acknowledges, it is not entirely clear from an empirical neuroscientific perspective precisely what role synapse selection plays in the neuroplastic changes associated with Substance Use Disorder. However, let us suppose that the relevant neural systems and sub-systems have indeed undergone dopamine-mediated synaptic selection such that, in accordance with Garson's analysis, these items have acquired the ontogenetic function of performing Substance Use Disorder.

If so then, in accordance with the principle that proper functions are proximal, we might conclude as follows: a complex system of neural traits spearheaded by the basal ganglia, the extended amygdala and the orbitofrontal/prefrontal cortex has acquired the proper ontogenetic neural function of Substance Use Disorder. This system is as such the trait T which has Substance Use Disorder as its ontogenetic function. Its constituent parts have as their proper ontogenetic function to contribute (however they in fact contribute) to the performance of this overall effect – assuming, as we do, that they have undergone relevant selection.

If so, then when constituent traits of this system – such as basal ganglia or the orbitofrontal cortex – contribute to the overall system-level effect of Substance Use Disorder, they are performing their ontogenetic function. This follows from the Generalised Theory and insofar Garson is correct. However, it does *not* follow from the Generalised Theory that these neural items are *therefore* not dysfunctional. This is because neural traits have evolutionary functions too.

Recall that the Generalised Theory encompasses all the evolutionary functions of the Traditional Theory, in addition to Garson's novel ontogenetic functions. As such, if Substance Use Disorder entails the failure of an evolutionarily selected effect of one of these implicated neural traits this would still be a dysfunction on the Generalised Theory – whether or not Substance Use Disorder results from synaptic selection processes. So does causing Substance Use Disorder entail failure of an evolutionary function in one of the implicated neural traits?

For the sake of simplicity, I consider only the role of the orbitofrontal cortex – a region of the prefrontal cortex located near the frontal lobes – in what follows. However, one could run the same argument with the amygdalae, the basal ganglia or the prefrontal cortex more generally (assuming that their contribution to Substance Use Disorder also infringes on their evolutionary functions).

The precise evolutionary function of the orbitofrontal cortex is naturally a matter of some controversy, but one theory, strongly supported by lesion studies, is that the orbitofrontal cortex has an important control function in certain types of decision-making processes. More specifically, the orbitofrontal cortex is hypothesised as playing a role in the suppression of inappropriate actions and impulses (Bechara & Van Der Linded, 2005; Ouellet et al., 2015). As such, if Substance Use Disorder interferes with this evolutionary function of the orbitofrontal cortex, then Substance Use Disorder still constitutes a dysfunction on Garson's view.

The notion that substance dependence adversely affects 'normal' human impulse control and response inhabitation is not only intuitive, it is borne out by the empirical evidence. Psychological research has shown associations between substance dependence and impairments in decision-making processes (Grant et al, 2000; Zhang et al, 2011). This finding as a matter of fact forms part of the evidence base for neuroscientific theorising around Substance Use Disorder and the role of the orbitofrontal cortex (Bolla, et al., 2003;

Torregrossa, et al., 2008). So Substance Use Disorder indeed interferes with an evolutionary function of the orbitofrontal cortex.

Accordingly, what follows from Garson's view is this: contributing to Substance Use Disorder (the effect of interest, E) is *both* a function *and* a dysfunction of the orbitofrontal cortex (the trait, T). Substance Use Disorder is both a case of biological function and a case of biological dysfunction. This finding is insufficient to reject the view that disorder entails dysfunction – as the Dysfunction Account states. Premise B is false, and the Dysfunction Account stands.

7.2. Premise A is false

As the demonstrated in the previous section of this paper, following Garson, contributing to Substance Use Disorder (E) is both a function and a dysfunction of the orbitofrontal cortex (T). In other words:

T is doing E
E is a (direct proper) function of T
E is a (direct proper) dysfunction of T

It seems then that Garson's Generalised Theory yields conflicting functional norms. The Generalised Theory implies that some single effect E of some single trait T can simultaneously be both functional and dysfunctional – in precisely the same sense (direct, proper, biological function).

Note first that this contradictory implication has no analogue on the Traditional Theory. The problem is not that some single trait T can have more than one function. The enlarged claw of the male fiddler crab has both a sexually selected aesthetic signalling function and a function as a weapon in confrontation with other males (Dennenmoser & Christy, 2013) both of which result from natural selection. Of course, the claw of some particular crab may be an effective signal, being large and ominous looking, without having any efficacy in a fight.

This however not what's going on with Substance Use Disorder. In the case of the fiddler crab there are two effects in play (E1, E2) – one which is being performed and another which is not. In the case of Substance Use Disorder, there is a single effect E which *is* being performed, and which is both a function and a dysfunction. A true analogue would be where a fiddler crab's claw *is* an effective weapon, and this is simultaneously both the fulfilment of an evolutionary function and the failure of an evolutionary function. And this simply cannot occur.

The conflicting functional norms yielded by the Generalised Theory arise because Garson is forcing two distinct sources of normativity to operate theoretically as one. It is a good indicator that ontogenetic and evolutionary functions are different in kind – analogous, perhaps, to how some single effect E can violate a social norm whilst fulfilling a biological one (think, for example, of teenage pregnancy). In what follows I argue that the contradictory implications of Garson's Generalised Theory for Substance Use Disorder (and moreover, the theoretical and empirical possibility of cases like it) suffices to reject the Generalised Theory outright.

Consider again Garson's parity of reasoning argument for the Generalised Theory. Garson argues that we must accept the Generalised Theory because parity of reasoning demands it:

> Consider why we accepted the traditional selected effects view. We did so because it made sense of three big puzzles of function … Since [the Generalised Theory] solves all the same problems, minus an arbitrary distinction, we should accept it.
>
> (Garson, 2019, p. 94)

There is however another type of function which accounts for the three big puzzles – that is, artefact functions. Artefact functions make sense of the function/artefact distinction, provide for dysfunctions and are explanatorily deep. Accordingly, any etiological theory generalising across ontogenetic, evolutionary and artefact functions would also seem to satisfy Garson's parity of reasoning argument – think for example of Larry Wright's original 1976 analysis which did just this.

So if parity of reasoning was really enough to motivate a *generalised* theory across types of function, Garson ought really commit himself to a generalised artefact-evolutionary-ontogenetic account of function.

However, he does not:

> I don't think of artifact functions and biological functions as two species of the same genus, like lions and tigers are two species of *panthera* … artifact and biological functions are different sorts of things…
>
> (p. 30, Garson, 2019)

It seems that when Garson stresses parity of reasoning, he is in fact leaning very heavily on the contention that we have *no principled reason* to distinguish evolutionary and ontogenetic functions. Accordingly, all that is needed to abandon Garson's generalised account of evolutionary and ontogenetic functions is some non-arbitrary reason to think that evolutionary and ontogenetic functions should not bundled together for theoretical purposes – just as Garson readily rejects a generalised theory across artefact and biological functions because he takes them to be different in kind.

And, of course, we have precisely such a reason. The Generalised Theory yields conflicting functional norms. This kind of normative contradictions is theoretically undesirable and good *prima facie* reason to think that the two types of function at play in the Generalised Theory (that is, evolutionary and ontogenetic functions) are (seeing as they disassociate) in fact different in kind. The contradiction in functional norms is however easily solved by simply abandoning Garson's commitment to generality about evolution and ontogenetic functions:

> T is doing E
> E is a (direct proper) *ontogenetic* function of T, and
> E is a (direct proper) *evolutionary* dysfunction of T

Garson's argument for the Generalised Theory fails. We have good principled, non-arbitrary grounds for preferring a disambiguated, specific theory of biological function. In other words, we should reject premise A.

References

Bolla, K. I., Eldreth, D. A., London, E. D., Kiehl, K. A., Mouratidis, M., Contoreggi, C. E. E. A., ... & Funderburk, F. R. (2003). Orbitofrontal cortex dysfunction in abstinent cocaine abusers performing a decision-making task. *Neuroimage*, *19*(3), 1085-1094.

Dennenmoser, S., & Christy, J. H. (2013). The design of a beautiful weapon: compensation for opposing sexual selection on a trait with two functions. *Evolution: International Journal of Organic Evolution*, *67*(4), 1181-1188.

Garson, J. (2019). *What biological functions are and why they matter*. Cambridge University Press.

Garson, J., & Papineau, D. (2019). Teleosemantics, selection and novel contents. *Biology & Philosophy*, *34*(3), 36.

Godfrey-Smith, P. (1994). A modern history theory of functions. *Noûs*, *28*(3), 344-362.

Grant, S., Contoreggi, C., & London, E. D. (2000). Drug abusers show impaired performance in a laboratory test of decision making. *Neuropsychologia*, *38*(8), 1180-1187.

Griffiths, P. E., & Matthewson, J. (2018). Evolution, dysfunction, and disease: A reappraisal. *The British Journal for the Philosophy of Science*, *69*(2), 301-327.

Koob, G. F., & Simon, E. J. (2009). The neurobiology of addiction: where we have been and where we are going. *Journal of drug issues*, *39*(1), 115-132.

Koob, G. F., & Volkow, N. D. (2016). Neurobiology of addiction: a neurocircuitry analysis. *The Lancet Psychiatry*, *3*(8), 760-773.

Lemoine, M. (2013). Defining disease beyond conceptual analysis: An analysis of conceptual analysis in philosophy of medicine. *Theoretical Medicine and Bioethics*, *34*(4), 309-325.

Millikan, R. G. (1989). In defense of proper functions. *Philosophy of science*, *56*(2), 288-302.

Neander, K. (1991). Functions as selected effects: The conceptual analyst's defense. *Philosophy of science*, *58*(2), 168-184.

Neander, K. (1995). Misrepresenting & malfunctioning. *Philosophical Studies*, *79*(2), 109-141.

Neander, K. L. (1983). *Abnormal Psychobiology: A Thesis on the 'anti-psychiatry Debate' and the Relationship Between Psychology and Biology* (Doctoral dissertation, La Trobe University).

Neander, K. (1998). Mental illness, concept of. Routledge Encyclopedia of Philosophy.

Schwartz, P. H. (2014). Reframing the disease debate and defending the biostatistical theory. *Journal of Medicine and Philosophy*, *39*(6), 572-589.

Torregrossa, M. M., Quinn, J. J., & Taylor, J. R. (2008). Impulsivity, compulsivity, and habit: the role of orbitofrontal cortex revisited. *Biological psychiatry*, *63*(3), 253-255.

Wakefield, J. C. (1992). The concept of mental disorder: on the boundary between biological facts and social values. *American Psychologist*, *47*(3), 373.

Wakefield, J. C. (2014). The biostatistical theory versus the harmful dysfunction analysis, part 1: is part-dysfunction a sufficient condition for medical disorder?. In *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine* (Vol. 39, No. 6, pp. 648-682). Journal of Medicine and Philosophy Inc.

Zhang, X. L., Shi, J., Zhao, L. Y., Sun, L. L., Wang, J., Wang, G. B., ... & Lu, L. (2011). Effects of stress on decision-making deficits in formerly heroin-dependent patients after different durations of abstinence. *American Journal of Psychiatry*, *168*(6), 610-616.