Biology & Philosophy (2022) 37: 30
https://doi.org/10.1007/s10539-022-09861-2

Against the generalised theory of function

Harriet Fagerberg[1,2]

Received: 17 September 2021 / Accepted: 9 June 2022 / Published online: 2 July 2022

Abstract
Justin Garson has recently advanced a Generalised Selected Effects Theory of biological proper function. According to Garson, his theory spells trouble for the Dysfunction Account of Disorder. This paper argues that Garson's critique of the Dysfunction Account from the Generalised Theory fails, and that we should reject the Generalised Theory outright. I first show that the Generalised Theory does not, as Garson asserts, imply that neurally selected disorders are not dysfunctional. Rather, it implies that they are *both* functional *and* dysfunctional. I argue on this basis that the Generalised Theory yields conflicting functional norms, and we that should reject it outright on these grounds.

Keywords Philosophy of biology · Philosophy of psychiatry · Neural selection · Biological function · Neuroplasticity

Introduction

This paper defends the Dysfunction Account of Disorder against Justin Garson's Generalised Selected Effects Theory of 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 theory. What in fact follows is that neurally selected disorders are *both* functional *and* dysfunctional. Establishing this clearly does not suffice to reject the thesis that disorder entails dysfunction, as the Dysfunction Account states. I proceed from here to argue that, given the fragility of Garson's 'parity of reasoning' argument in support of the Generalised Theory, this inconsistent consequence of his theory is sufficient grounds to reject it outright.

✉ Harriet Fagerberg
harriet.fagerberg@kcl.ac.uk

[1] King's College London, London, UK

[2] Humboldt-Universität zu Berlin, Berlin, Germany

The traditional theory

In his 2019 book *What biological functions are and why they matter,* Garson contrasts his novel '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)[1]

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.

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 succeeds in also encompassing a range of developmentally selected functions by appealing to effects selected by differential retention and reproduction through ontogeny, as well as by differential reproduction through evolution. (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.)

[1] Exact formulations of the Traditional Theory actually vary significantly. For example, Millikan requires some form of reproductive mechanism, while Neander emphasises selection at the level of the genotype (Millikan 1989; Neander 1991). The important point for our present purposes it that the Traditional Theory excludes attributing direct proper biological functions to most ontogenetic functions (see Garson 2019, for a supporting argument). My definition above merely seeks to accentuate this distinction.

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.[23] 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's account, the surviving synapses acquire through this process the function of doing whatever it was that caused them to be selectively retained relative to the others through ontogeny—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's argument for the generalised theory

Garson contends that the same argument which supports the Traditional Theory also supports his Generalised Theory. According to Garson, philosophers of biology 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 the capacity of the nose to hold up spectacles. Because the nose *evolved* to smell (and not to hold up spectacles) that is its function. (2) The Traditional Theory also successfully accounts for the apparent 'normativity' of functions—that is, the theoretical possibility of a trait failing to function as it 'should'. Because a trait's function is determined by its history of selection (as opposed to any of its current activities or characteristics) it can fail to perform a function that it is 'supposed' to perform. Thus, the Traditional Theory makes sense of the perplexing normative force of functional statements. (3) Finally, we sometimes cite functions as explanations for the structure and existence of traits. "Why do Zebras have stripes? For camouflage." The Traditional Theory tells us that a function is an effect which *caused* its corresponding trait to be naturally selected—in this sense, the effect really *does* explain the trait. 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)

[2] For the those unfamiliar with the function literature, *proper* functions are functions in the sense of 'function of' rather than 'functions as'. Proper functions are often contrasted with causal role functions (Cummins 1975). Many philosophers are 'pluralists' in that they accept both of these senses of function are legitimate (see e.g. Garson 2018; Godfrey-Smith 1993).

[3] *Direct* proper functions have their functions inherently. *Derived* proper functions only have their functions in virtue of the direct proper function of the mechanisms which produce the derived proper functions (for more on this see Millikan 1989, 1999).

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

So what does it mean to accept a 'generalised' theory of function, as Garson contends that we must? It is not simply the claim that applying the term 'function' to ontogenetic functions would be semantically correct. Philosophers and biologists already do this in various contexts, so this would be kicking at an open door. Rather, Garson is committed to something like the following: evolutionary and ontogenetic selection processes yield functions of *the same kind*.[4] Ontogenetic and evolutionary functions play—or should play—the same theoretical role. Restricting one's concept of direct proper function to include only evolutionary functions, even in specific domains, is theoretically unprincipled.[5]

Garson proceed to argue that acceptance of the Generalised Theory spells trouble for the Dysfunction Account of Disorder.

The dysfunction account

The 'Dysfunction Account of Disorder' (from here, the 'Dysfunction Account') 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)[6]

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

[4] This can be inferred from his contention that a Generalised Theory of function across biological and artefact functions would be inappropriate because biological and artefact functions are different in kind (see Garson 2019 p. 22).

[5] Does it have to be this strong? It does, and *why* it does will become clear as we go on If we were justified in disambiguating or discounting the Generalised Theory of function for specific theoretical purposes, such as that of defining medical disorder, then the detrimental impacts which Garson infers for the Dysfunction Account simply wouldn't follow.

[6] The Dysfunction Account of Disorder is often associated with Jerome Wakefield and his so-called 'harmful dysfunction' analysis of mental disorder but is perhaps more accurately attributed to Karen Neander who advocated this view as early as in her unpublished 1983 doctoral thesis (Wakefield 1992; Neander 1983). The Dysfunction Account also owes a debt to Christopher Boorse's 'biostatistical theory' of disease, however Boorse vigorously rejects the selected effects theory of function and appeals instead to statistically typical contributions to goals (1976, 1977).

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 novel challenge to the Dysfunction Account which explicitly proceeds from the Generalised Theory will take a similar form.[7]

(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 possess direct proper functions, thereby increasing the likelihood that a given condition is functional, not dysfunctional." (Garson 2019, p. 178). 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.
>
> (Garson 2019p. 180; see also Garson and Papineau 2019)

(C) From A and B, we should accept that Substance Use Disorder is functional and not dysfunctional.[8]
(D) Substance Use Disorder is a case of mental disorder.

[7] Garson also contends that the Dysfunction Account cannot account for cases of developmental mismatch—that is, when a mechanism is performing its evolutionary function of yielding a certain phenotype given certain signals or circumstances, but that phenotype turns out to be disadvantageous to the organism (due for example to a change in its environment) (Garson 2014, 2019, 2021; see also Matthewson and Griffiths 2017). This case is equally problematic on the Traditional Theory, and I will restrict my scope here to objections that apply to the Generalised Theory specifically.

[8] An anonymous reviewer pressed me on whether it is appropriate to speak of Substance Use Disorder as being a function or a dysfunction. Perhaps Substance Use Disorder is a broader phenomenon than its neurobiological cause and depends on contextual and social factors. This depends, I think, on how we view the metaphysics of Substance Use Disorder. Can Substance Use Disorder be theoretically defined in terms of its neurobiological cause or not? I won't rule on this here. However, my argument can alternatively be read as pertaining to whether Substance Use Disorder *involves* dysfunction or *is caused by* dysfunction.

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.[9]

(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.

Rebutting Garson's critique

In what follows, I will show that Garson's premise B is false—it does not follow from the Generalised Theory that Substance Use Disorder is functional and therefore not dysfunctional. I will then argue that premise B being false should also lead 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.

We should reject premise B

If we accept the Generalised Theory and if we accept that Substance Use Disorder is the result of synapse selection of its underlying neural trait, then Garson is correct—Substance Use Disorder is functional on his view.

However, it does *not* follow from this alone that Substance Use Disorder is *not* dysfunctional. Garson neglects to consider that traits of the brain have evolutionarily selected effects too (evolutionary functions), and that the Generalised Theory encompasses all of these proper functions as well. As such, determining that Substance Use Disorder is synaptically selected and therefore functional on the Generalised Theory is plainly insufficient to establish that it is *not dys*functional—it remains open that Substance Use Disorder constitutes an evolutionary dysfunction.

If Substance Use Disorder is an evolutionary dysfunction, then Garson is *wrong* to conclude that Substance Use Disorder is *not* dysfunctional on the Generalised Theory, and his critique of the Dysfunction Account accordingly fails. In what follows I shall argue that Substance Use Disorder *is* indeed an evolutionary dysfunction—quite independently of whether it came about via synapse selection.

In order to establish whether Substance Use Disorder constitutes an evolutionary dysfunction, we must first get clear on precisely which neural trait has the function of Substance Use Disorder on Garson's view. 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' (see also

[9] I grant Garson this for present purposes, but there is considerable debate in the literature as to whether inclusion in official classification should be taken as conclusive (see e.g. Schwartz 2014; Lemoine 2013; Griffiths and Matthewson 2018; Cooper 2020).

Neander 1995). For example, it is the proper, function of the human 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" (2019, p. 119).

It follows that the trait T which has the function Substance Use Disorder must be the overall neural system which in fact yields this effect. Unless some particular synapse is performing Substance Use Disorder all on its own, which is unlikely, it would be fallacious to attribute to it a system-level effect like Substance Use Disorder.[10] The question thus becomes, what does the empirical evidence indicate about which neural trait in fact performs Substance Use Disorder in the addicted brain?

The trait which in fact performs Substance Use Disorder in the brain 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 abuse. According to one neuropsychological theory, 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 and Simon 2009). This is supported by functional neuroimaging studies showing that substance dependence is associated with neuroplastic changes in the basal ganglia, the extended amygdala and the orbitofrontal/prefrontal cortex (Koob and Simon 2009; Koob and Volkow 2016).

As Garson acknowledges, it is not entirely clear from the evidence precisely what role synaptic selection plays in these neuroplastic changes. However, let us suppose, in accordance with Garson's account, that this neural system has undergone dopamine-mediated selection and thus has acquired the overall ontogenetic function of yielding Substance Use Disorder. If so, then the trait T which has as its proper function to perform Substance Use Disorder is something like a complex neurobiological system spearheaded by the basal ganglia, the extended amygdala and the orbitofrontal/prefrontal cortex. Its constituent parts, such as the neuroplastically adapted orbitofrontal cortex, thus have as their proper ontogenetic function to contribute, howsoever they in fact contribute, to this overall system-level effect—just as the heart contributes to circulation by beating.

We are now in a position ask: does Substance Use Disorder infringe upon the evolutionary functions of any of the implicated neural traits, such that Substance Use Disorder would have to be considered a dysfunction of these traits as well? In order to answer this question, we need to know what the evolved functions of the relevant traits are. For the sake of simplicity, I will consider only the orbitofrontal cortex—a region of the prefrontal cortex located near the frontal lobes—in what follows.[11]

[10] I will unpack this more thoroughly in "What has the function? Levels, traits and proximate effects".

[11] However, in principle, one could run the same argument with the amygdalae, the basal ganglia or the prefrontal cortex as a whole, assuming that their contribution to Substance Use Disorder likewise infringe on their evolutionary functions.

The precise evolutionary function of the orbitofrontal cortex is naturally a matter of some controversy within the relevant sciences.[12] However, one influential theory, strongly supported by lesion studies, is that the orbitofrontal cortex has characteristic control function in certain types of decision-making processes. More specifically, the orbitofrontal cortex is hypothesised as playing an important role in suppressing or inhibiting inappropriate actions, behaviours and impulses: "[S]ubjects with [orbitofrontal cortex] lesions are more impulsive overall compared both to normal controls and to those with non-[orbitofrontal cortex] brain damage … [R]ecent functional neuroimaging studies have reported that escalating risk-taking behaviour in healthy volunteers is associated with decreased activity in the ventromedial prefrontal cortex (of which the [orbitofrontal cortex] is a key component)" (Ouellet et al. 2015, p. 28; see also Bechara and Van Der Linded 2005). Suppose this theory is correct. If so, then Substance Use Disorder *would* constitute an evolutionary dysfunction *if* it interfered with the orbitofrontal cortex's evolved decision-making capacities.

One might be inclined to grant that Substance Use Disorder interferes with healthy decision making and inhibitory control on purely intuitive grounds. From a folk-psychological perspective, pathological drug dependence does seem to be characterised by an inability to inhibit inappropriate actions and impulses—exemplified most obviously by the inability to supress the impulse to take drugs. We need not rely on our intuitions here however as a variety of neuropsychological evidence also lends credence to this hypothesis (Grant et al. 2000; Zhang et al. 2011). For example, Ekhtari et al. found in a 2010 review "substantial evidence that altered decision-making in people with [Substance Use Disorder] is not limited to drugs but extends to other decision-making situations." (Ekhtiari et al. 2017, p. 26). Koffarnus and Kaplan similarly note that it is "increasingly clear that decision-making processes are dysfunctional in addiction and that this dysfunction may be fundamental to the initiation and maintenance of addictive behavior" (Koffarnus and Kaplan 2018, p. 71).

Indeed, psychological findings of decision-making dysfunction in individuals with Substance Use Disorder in fact form part of the evidence base for neuroscientific theorising around addiction and the contributory role of the orbitofrontal cortex: "[T]he inability of humans who are cocaine-addicted … to make adaptive decisions in the present, and their corresponding difficulty in learning from unexpected or changing outcomes to improve decision-making in the future [may result in part] from long-lasting cocaine induced neuroadaptations in the orbitofrontal cortex [and other] interconnected brain regions" (Lucantonio et al. 2012, p. 359; see also Bolla et al. 2003; Torregrossa et al. 2008; Schoenbaum and Shaham 2008). In other words, impairments of healthy decisions making processes via neuroplastic changes to the orbitofrontal cortex is *part* of Substance Use Disorder.

[12] Indeed, it is rarely discussed in these terms. However, talk of '*the* function' of the orbitofrontal cortex in neuroscientific and neurological contexts can generally be assumed to be employing at a notion of evolutionary proper function (or some other notion of proper function with roughly the same extent) (see Neander 2017 for a supporting argument).

Let us recap. If the empirical premises of my argument are true, then we can conclude as follows.[13] The orbitofrontal cortex (the trait, T) contributes to the performance of Substance Use Disorder (an effect, E). T has undergone synapse selection such that, when T contributes to E, T is (per the Generalised Theory of function) performing its direct proper function. However, because T has an evolved decision-making function as well, and because yielding E *infringes* on the evolved decision-making capacities of T, T causing E is *also* a *dys*function on the Generalised Theory. In other words:

T is doing E.
E is a direct, proper biological function of T.
E is a direct, proper biological dysfunction of T.

In sum, Garson fails to establish that Substance Use Disorder is functional and *therefore* not dysfunctional—this does not follow from his view. At best, he demonstrates that it is *both* functional *and* dysfunctional, and this outcome is quite clearly insufficient to reject the hypothesis that disorder entails dysfunction (as the Dysfunction Account states). Premise B is false, and the Dysfunction Account stands.

We should reject premise A

In virtue of its generality across evolutionary and ontogenetic functions, Garson's theory has a puzzling implication: a single effect E of a single trait T can simultaneously be *both* functional *and* dysfunctional. In what follows, I shall first motivate why conflicting function-dysfunction attributions is some cause for concern. I will then argue that the possibility of such conflicts should lead us to reject Garson's parity of reasoning argument in support of the Generalised Theory.

Conflicting attributions of function and dysfunction

Why worry about the possibility of conflicts? So what if Garson's theory has this slightly odd implication?

Note first that these conflicts have no analogue on an ungeneralised theory of biological function which restrict its scope to evolutionarily selected effects. The problem is not, for example, that some single trait T can have more than one biological function. This can occur on the Traditional Theory as well, and is not a cause for concern. For example, 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 and Christy 2013). Of course, the claw of some particular crab may be an effective signal without having any efficacy in a fight—suppose it is large and ominous looking but lacks grip

[13] These assumptions are not unquestionable (see e.g. Stalnaker et al. 2015) and, of course, establishing their truth is not a matter of philosophical inference. They are empirical best ultimately to be settled by the relevant sciences.

strength. In this sense, trait T can simultaneously be *both* functional and dysfunctional also on the Traditional Theory.

This is no analogue to Substance Use Disorder, however. In the case of the male fiddler crab, there are *two* selected effects in play (E1, E2)—one which the trait *is* performing, and another which the trait *is not*. In the case of Substance Use Disorder, there is a single effect E (contributing to Substance abuse Disorder via changes to the orbitofrontal cortex) which *is* being performed, and which is *both* a function *and* a dysfunction. The decision-making dysfunction is not separate from Substance Use Disorder, but part of it—it is *how* the orbitofrontal cortex contributes to addiction. In the case of the fiddler crab claw, conversely, there is one effect E1 which is being yielded as designed (appearing impressive) and another effect E2 *not* being yielded (having efficacy in a fight).

Conflicting function-dysfunction attributions of the sort yielded by the Generalised Theory in the case of Substance Use Disorder are only possible because Garson is forcing distinct sources of normativity to operate as if they were one. It is a strong indication that evolutionary and ontogenetic functions are not of the same *kind.* It is analogous to the dissociations which can obtain between biological functions, on the one hand, and artefact functions or social functions on the other. Consider how a single effect E can constitute an (artefact) function whilst simultaneously constituting an (evolutionary) dysfunction. A pair of bound feet in nineteenth century China fulfilled their artefactual function of appearing dainty and attractive (per to the conventions of their day) whilst, simultaneously and by the very same token, being unable to function as designed by evolution (being unable to walk). Or consider how teenage pregnancy is perfectly functional from an evolutionary perspective and yet fails to conform to prevailing social norms (see also Neander 1995). I take it that ontogenetic and evolutionary functions yield conflicts of the same sort.

In insisting that evolutionary and ontogenetic functions be subsumed under a single theoretical heading, Garson forfeits the resources to recognise this distinction in kind and thus leaves us in the puzzling position of having to conclude that Substance Use Disorder is *both* a function *and* a dysfunction—in precisely the same sense of function. Accordingly, the implications of the Generalised Theory for any philosophical account which relies either on the concept of function or the concept of dysfunction—such as the Dysfunction Account in philosophy of medicine or the teleosemantic theory of representation—are simply ambiguous. The Generalised Theory does not falsify the Dysfunction Account; it renders it obscure and equivocal, and hard to apply in a consistent and useful way.

This, I take it, is a theoretical defect. However, all I need to establish for now is the following, weaker premise:

> The possibility of conflicting function-dysfunction attributions constitutes a non-arbitrary and principled reason to favour a specific (that is, ungeneralised) theory of biological function.[14]

Given the fragility of Garson's parity of reasoning argument, this weaker premise suffices to reject the Generalised Theory outright.

Rebutting the parity of reasoning argument

Let us now return to Garson's parity of reasoning argument. Garson contends we must accept the Generalised Theory in place of the traditional theory "on pain of inconsistency" (Garson 2019, p. 93). Further inspection reveals that there are two possible interpretations of the parity of reasoning argument. On the first interpretation, Garson's position on artefact functions undermines him. The second interpretation rests on an unsubstantiated, unargued premise that is easily rejected.

The first interpretation of the parity of reasoning argument goes something like this:

(1) Because evolutionarily selected effects are normative, explanatorily deep and distinguishable from accidents, evolution by genetic selection is a function bestowing process.
(2) Ontogenetically selected effects are normative, explanatorily deep and distinguishable from accidents.
(3) By parity of reasoning, ontogenetic selection is a function bestowing process.
(4) So, ontogenetically selected effects are *functions.*

This appears to be the structure of Garson's inference in some, but not all, formulations of his parity of reasoning argument. For example, while discussing Millikanian derived proper functions he concludes as follows: "The reason why we identify functions with selected effects [is] because doing so solved all the big problems of biological function. By parity of reasoning, any process that can be used to solve those puzzles should also give rise to new functions, too" (2019, p. 73) Similarly: "I agree that we should be able to say that trial-and-error … creates new functions. After all, it makes sense of the function/accident distinction, and the explanatory and normative side of functions" (2019, p. 74).

Let us assess this argument. If we accept premise 1 and 2, then yes: we must accept, on pain of inconsistency, that ontogenetically selected effects are functions.[15] This clearly underdetermines however that we should *accept the Generalised*

[14] By a 'specific' or 'ungeneralised' theory of function I mean a theory which recognises the distinction in kind between evolutionary and ontogenetic functions.

[15] It is not clear to me that premise 1 is true however—is this really how people reason about functions? It seems to me, rather, that people generally take it as a given that natural selection is function bestowing, and then simply *observe* that the functions yielded by natural selection are normative, explanatorily deep and distinguishable from accidents—no inference required.

Theory. It does not follow from the fact that ontogenetically selected effects, like evolutionarily selected effects, are functions of *some sort*, that ontogenetic and evolutionary effects are functions of precisely the *same* sort and thus must be accounted for by a single, general theory. We could readily accept that ontogenetically selected effects count as functions, without making the further leap of demanding a general, unified theory of the kind Garson defends. Suppose we have reason to suspect, for example, that ontogenetic and evolutionary functions are different in kind, fear the possibility of conflicting functional norms, and therefore prefer not to yoke evolutionary and ontogenetic functions together in a single general theory. Nothing about Garson's parity of reasoning argument prohibits us from adopting such a position.

Indeed, Garson *must* accept the legitimacy of this midway position—on pain of inconsistency—because it is precisely analogous to his own position on artefact functions. Artefacts can become dysfunctional (normativity), we sometimes speak as if the functions of artefacts *explain* their features (explanatory depth), and artefact functions are distinguishable from accidental effects.[16] If the parity of reasoning argument as presented in the above *mandated* the acceptance of a generalised theory, then Garson would have to accept a general theory across artefact and biological functions. Yet 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
>
> (Garson 2019, p. 30)

Garson accepts that artefact functions are functions, in a sense, but denies the further, stronger claim that they are functions in the *same* sense as biological proper functions. In precisely the same way, I accept that ontogenetic selection is function bestowing but deny the appropriateness of a *generalised* account. Garson's parity of reasoning argument, thus understood, fails to sufficiently motivate the Generalised Theory.

The first interpretation of the parity of reasoning argument thus fails. There is however a second possible interpretation of Garson's parity of reasoning argument. It goes something like this:

(1) We accepted the Traditional Theory of function because of its ability to account for the three puzzles—the normativity of functions, the explanatory depth of functions, and the function/accident distinction.
(2) The Generalised Theory can account for the three puzzles.
(3) From 1 and 2, we have *equally* good reason to accept the Generalised Theory as the Traditional Theory.
(4) Distinguishing evolutionary and ontogenetic functions theoretically would be arbitrary, unprincipled and/or pointless.
(5) So, we should accept the Generalised Theory over the Traditional Theory.

[16] I elaborate on this in "The explanatory depth of artefact functions".

This argument is different than that outlined in the above. The first interpretation of Garson's parity of reasoning argument concerns whether ontogenetic selection processes count as function-bestowing processes, which, as noted, underdetermines whether we should accept a general theory of evolutionary and ontogenetic functions—just as agreeing that intentional design is a function bestowing process underdetermines whether artefact and evolutionary functions should be accounted for by a single, uniform theory. This second interpretation of the argument however concerns our reasons for accepting any theory of function, and our reasons for preferring the Generalised Theory in particular. This seems to be more the structure of Garson's argument in some places:

> 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)

Elsewhere:

> My main argument for the [the Generalised Theory] is that is solves all the puzzles of function, without pointless restrictions.
>
> (Garson 2019, p. 101)

According to this version of the parity of reasoning argument, each theory's ability to account for the three big puzzles of biological proper function would appear to give us *equally* good reason to accept the Generalised Theory as the Traditional Theory (assuming, as Garson does, that making sense of the three puzzles is the *only* desideratum for a theory of function, see also Garson and Papineau 2019). So how do we choose between them? Garson's answer is that we have to choose the Generalised Theory because an ungeneralised account of function would be 'arbitrary' and 'pointless'.

As now becomes clear, Garson is leaning very heavily on the unargued, unsupported contention that favouring an ungeneralised account of biological function over the Generalised Theory would be arbitrary and unprincipled. This renders his overall argument weak and easily rejected. All we would need to reject the Generalised Theory outright is some non-arbitrary reason to prefer a disambiguated, ungeneralised account of function (in place of the Generalised Theory). Do we have such a reason?

Of course, we have precisely such a non-arbitrary, principled reason on hand: the possibility of contradictory function-dysfunction attributions resulting from conflicting functional norms within the Generalised Theory. It is a perfectly principled and legitimate to prefer an account that does *not* obscure the theoretical distinction between evolutionary and ontogenetic functions – as noted, this sort of ambiguity threatens to undermine the application of concept of function in a range of philosophical contexts. Garson cannot sensibly maintain that preferring

a theory of function which yields unambiguous and consistent attributions of function and dysfunction is arbitrary and unprincipled.

In other words, we should reject premise 4 of the parity of reasoning argument as outlined in the above. If we reject premise 4, then we are free to reject the parity of reasoning argument. If we reject the parity of reasoning argument, then we are free to reject the Generalised Theory (premise A).

In short; whether we read Garson's argument for the Generalised Theory as pertaining to what processes are rightly construed as function bestowing, or which of our theories rightly merit acceptance, Garson's parity of reasoning argument fails to motivate acceptance of the Generalised Theory.

Objections

I have argued as follows. The Generalised Theory does not imply that Substance Use Disorder is functional and *therefore* not dysfunctional; rather, it implies that Substance Use Disorder is functional *and* dysfunctional. The possibility of such conflicts—given the fragility of Garson's parity of reasoning argument (however interpreted)—suffices to reject the Generalised Theory outright.

In the following section, I respond to three possible lines of retort. In "The Explanatory Depth Of Artefact Functions", I consider whether artefact functions are really explanatorily deep. Then, in "What has the function? Levels, traits and proximate effects", I consider whether it is really the complex neural trait that bears the ontogenetic function Substance Use Disorder on Garson's view or whether it is individual synapses within this trait. In "Autoantibodies, cancer and feral children: dispelling residual doubt", I respond to concerns about the exclusive focus on Substance Use Disorder by outlining three other (possible) cases of conflict.

The explanatory depth of artefact functions

I have said that artefact functions are explanatorily deep because we speak as if their effects explain their existence. However, perhaps Garson requires more for explanatory depth. Perhaps Garson requires that it be *literally true* that an *actual* effect of the trait is explanatory. If so, then Garson is not inconsistent when he rejects a general theory of biological and artefact functions for artefact functions are not explained by their *actual* effects but by their *intended* effects. Hence, they are not explanatorily deep.

I shall briefly comment on why I think this interpretation of Garson's view is incongruent with Garson's own statements on the matter before showing that, in any case, it would not make sense of Garson's position.

The way Garson initially lays out 'explanatory depth', it is a prima facie feature of functions which an adequate theory thereof needs to make sense of. He writes: "The idea that functions are explanations is not some philosophers invention. It is a robust feature of how scientists and lay people alike think and talk about them" (2019, pp. 13–14). Garson gives the example of the Zebra. In attempting to determine the

function of its stripes, we might ask: 'Why do Zebras have stripes?' In response, we might entertain a number of hypotheses: Zebras have stripes for camouflage, or because the stripes deter flies.

Undeniably, we think and talk of artefact functions in this way too. Imagine that we unearth an historical artefact and we're not sure what it is. We note that it has a hook attached to it. In seeking to determine the function of this artefactual feature, we might ask each other: "Why does this artefact have a hook?" In response, we entertain a number of functional hypotheses: it has a hook for hanging on the wall, or because it looks cool. That is, we speak as if the functions of artefacts *explain* their features. Thus, artefact functions display the same prima facie explanatory depth as biological functions.

However, let us suppose that this is not what Garson meant after all. Rather, explanatory depth is a substantial, theoretical feature of functions which we need a philosophical theory to identify correctly. Explanatory depth is *only* present where trait T is explained by its *actual* effects. What biological functions have in common, in contrast to artefact functions, is that "an *actual* effect of the trait explains the trait's existence, not merely a presumed or desired effect" (2019, p. 30, emphasis mine).

The problem for Garson is that it is *not* true that biological (retained and reproduced) functions have this feature in common in contrast to artefacts. In fact, it is only selectively retained traits (such as Garson's selectively retained synapses) which are explained by *their actual* effects. Evolved, reproduced traits, like the human heart, are explained—not by *their actual* effects—but by past effects of traits of their type (see Neander and Rosenberg 2012; Nanay 2010). My particular heart was never selected for any actual beating that it does in my body. Selection never acted on my particular heart. Instead, natural selection acted on the hearts of my ancestors, and it is in virtue of *this* selection that my heart has a function.

As such, if the relevant trait being explained by *its actual* effects (rather than intended effects or past effects of the trait-type) is what marks a function as *truly* explanatorily deep and uniform in kind, then Garson should conclude that reproduced functions are *not really* explanatorily deep and, moreover, distinct in kind from retained functions.

In short, there are two ways to read Garson on explanatory depth, but neither renders his position on artefact functions coherent. One way or another, Garson is inconsistent in his insistence that evolutionary and ontogenetic functions are of the same kind and must be accounted for by a general theory.

What has the function? Levels, traits and proximate effects

I have assumed that it is the overall neural trait which actually performs pathological substance use (i.e. a complex of neuroplastically adapted sub-systems of the basal ganglia, the extended amygdala and the prefrontal, including the orbitofrontal, cortex) that has been ontogenetically selected for the performance of substance abuse through synapse selection. One might object to this characterisation in the following way. Garson is clear that mere retention (that some synapse, circuit or system is

retained because it yields an effect) is not sufficient for a new ontogenetic function to arise.

> [T]hat would be devastating for my theory of function. For it would mean that virtually any activity-dependent synapse change would count as "selection" and hence could create new functions. Panic attacks would create new functions, simply because they strengthen some synapses and not others … [T]he differential retention of synapses must occur within the same population. … In synapse selection, these interactions are competitive. One synapse is eliminated because another is retained.
>
> (Garson 2019, pp. 96–97)

A population, according to Garson, is a group within which there is a high degree of fitness relevant interactions—especially, competition. This requirement is satisfied in the case of synapses because there is evidence of competition (Garson 2019). One might object, however, that the complex neural trait I identified in 7.1. is *not* part of a population of fitness interacting traits and, therefore, even if it is retained, and synapse selection is causative in this retention, it does not count as bearing an ontogenetic function.

Instead, *individual synapses* within this complex trait bear the novel ontogenetic functions because *they* have been selectively retained within a population of individual competing synapses. If it is *individual synapses*, rather than the more complex neural system, which bear the novel ontogenetic function aren't there really *two* traits in play (T1, T2)? If so, is there *really* a conflict of the sort I allege?

In what follows I shall argue that the assumptions I made regarding the interpretation of Garson's position in "We should reject premise B" were charitable, and that dropping to the level of synapses yields problems for Garson which threaten his theory and his critique of the Dysfunction Account (albeit, via another route).

Remember that Garson needs to hold that there is a disorder—in this case, Substance Use Disorder—which is functional on the Generalised Theory.[17] I assumed in the above that, somehow, the overall neural trait which in fact performs substance dependence has acquired the retained, ontogenetic function of Substance Use Disorder via synapse selection. As I went on to argue, this yields a conflict—this complex trait does not *just* have an ontogenetic function, but evolutionary functions too.

Suppose that we adopt the alternative interpretation of Garson's commitments as outlined in the above. On this view, it is *individual synapses* that bear Garson's novel ontogenetic functions—not complex neural traits. How might we specify the *proximate* ontogenetic function of a single retained synapse? It is not clear what to say about this. What is clear, however, is that no single synapse bears the proximate function of a complex behavioural effect like pathological drug dependence. Individual synapses contribute to this effect in very distal terms, but we cannot attribute a distal effect to a *single* synapse within the larger system that yields it. That would

[17] Otherwise, he would fail to establish, as he must if he hopes to refute the Dysfunction Account, that there is something which *is* a disorder but is *not* dysfunctional.

be, as Garson puts it, 'a fallacy of division'; and, indeed, a much more egregious fallacy that the one he considers and rejects—akin, perhaps, to attributing the overall function of the cardiovascular system to a single cell.

Perhaps Garson would interject that it is not *individual synapses* which bear the function of Substance Use Disorder, but a *group* of synapses. A group of synapses, within the complex neural trait which in fact performs Substance Use Disorder has as its ontogenetic function to cause Substance Use Disorder, and it *is* performing this function whilst the complex neural system *isn't.* Again, two traits in play and no conflict.

Firstly, note it is not clear that this is compatible with proper functions being proximal either—Substance Use Disorder is arguably more complicated still. It is highly likely that other neuroplastic mechanisms, such as retention caused by neurons 'firing together and wiring together', is involved in Substance Use Disorder as well (Hogarth et al. 2013). Can we really attribute this overall effect to a specific group of retained synapses? Perhaps not. My hunch is that the closer you get to a group of synapses which could plausibly bear the proximate function of Substance Use Disorder the closer you get to the neural system outlined in 7.1. We shall put this worry aside for now however as there is a more fundamental problem with the move from synapse selection to *groups* of synapses selection.

Garson has argued persuasively that synapses compete with other synapses for differential retention (Garson 2019; see also Garson 2017). He has not shown, however, that *groups* of synapses compete with *groups* of synapses for differential retention. By analogy, it does not follow from the fact that selection acts on *organisms* within a *population of organisms*, that selection acts on *groups* of organisms within a *population of groups* of organisms. (Indeed, group selection is highly controversial in evolutionary theory.)

In pursuing this line, Garson is moving from one level of selection (selection between *synapses*) to another (selection between *groups* of synapses). This move cannot be made by mere extension but requires additional argument and evidence—just as we cannot simply *infer* group selection from selection between organisms, we cannot infer *group synapse selection* from selection between *synapses.* The former simply does not follow from the latter. What reason do we have to think that groups of synapses compete with other groups of synapses for differential retention such that they count as entities which bear ontogenetic functions? As it stands, Garson's account is silent on this point.[18]

[18] Is it possible that groups of synapses compete with other groups of synapses for retention? Perhaps, but Garson does not explicitly make this case. He discusses the possibility of group selection of neurons but, as he himself notes, it lacks empirical support. In the realm of synapses, he consistently speaks as if selection acts on the individual synapse: "Synapse selection takes place when two or more neurons synapse onto the same target, for example, another neuron or even a muscle fiber. These synapses behave differently (say, one of them is more active than the other). Because of these differences, one synapse is retained and the other eliminated. Crucially, these two events (the retention of one and elimination of another) are not causally independent. Instead, there is a competitive process that takes place between them. One is eliminated because the other is retained" (2017, p. 532). Moving from the selective retention of synapses to the selective retention of groups of synapses requires additional argument. And it is up to Garson to make this case convincingly.

If *groups* of synapses do not satisfy conditions for being selectively retained traits within a population of synapse-groups then *groups* of synapses are not the sort of things that have ontogenetic functions *at all* on Garson's view. If groups of synapses do not have ontogenetic functions at all, then *no* group of synapses has Substance Use Disorder as a function either. Worse yet, if complex neutral traits such as constellations of synapses or neuroplastically adapted brain systems do not count as having ontogenetic functions on Garson's view, then it is hard to see how the Generalised Theory could yield *any* interesting retained neural functions, as synapses rarely do anything interesting on their own. Thus, the utility and application of the Generalised Theory is called into question.[19]

In sum, if it is the *overall complex neural trait* which bears the ontogenetic function, then there is a conflict. If it is *individual synapses* that bear Garson's ontogenetic functions, then none of them have Substance Use Disorder as a proximate function, and thus Garson fails to establish that Substance Use Disorder is functional. If it is *groups* of synapses that have Substance Use Disorder as a function, then Garson needs to make this case more convincingly as it is not clear that *groups* of synapses count as the sorts of things that have functions *at all* on Garson's view. In short, either there *is* a conflict and my argument stands, or Garson fails to establish that Substance Use Disorder is functional on the Generalised Theory (premise B) and, thus, his critique of the Dysfunction Account fails.

Autoantibodies, cancer and feral children: dispelling residual doubt

Suppose that Garson is *still* not convinced by the case of Substance use Disorder. Remember then that the second part of my argument (outlined in We should reject premise A) does not rely on this case in particular. I focused on Substance Use Disorder in order to neutralise it as a counter example to the Dysfunction Account (which I sought to defend), however it is not necessary that this *particular* case turns out to be a conflict for the rest of my argument to follow. All that is required is that these sorts of conflicts can, in principle, occur on the Generalised Theory. As such, any analogous case will do just as well.

In what follows, I outline three sets of possible cases of some single effect E of some single trait T being functional (in ontogenetic terms) and dysfunctional (in evolutionary terms) at the very same time and for the very same reason. (Scope does not permit me to assess these in the same depth as Substance Use Disorder, but they should give an indication as the range of possibilities.) Finally, I show that conflicts are not exhausted by these sorts of cases; they could also occur in the reverse.

[19] I will not pursue this argument any further here, but it is for Garson to clarify and refine his account in this regard. My hunch is that the sort of refinement needed can only be attained by first giving up on his commitment to a general theory.

Autoantibodies

In his 2019 book, Garson considers another case of ontogenetic selection which, he argues, creates new biological functions: antibody selection.

> At birth, a mechanism of genetic recombination generates billions of genetically distinct B cells in the bone marrow. Each cell is coated with antibodies … when an antibody meets a foreign particle with a similar shape (the antigen), its corresponding B cell is massively replicated. … After a B cell has been massively replicated it undergoes a second round of selection called "affinity maturation". … [A]fter the antibody meets its antigen, its B cell starts replicating, but this replication process is unfaithful. That's because B cells are equipped with an enzyme, activation induced deaminase (AID), whose job it is to litter the B cells' genetic code with mutations ("somatic hypermutation"). When a B cell replicates, it creates a batch of daughter cells that genetically differ from each other and from their parent.
>
> (Garson 2019, pp. 69–70)

Through this process, the body develops antibodies that have a high degree of affinity for a particular foreign particle. Particular antibodies that result from selective replication come to have an ontogenetic function—their function is doing whatever caused their type to be replicated (i.e. binding to this particle).

Suppose that the evolved function of any antibody—as a result of genetic selection—is to bind to foreign particles or non-self antigens.[20] When an antibody has been selected for binding to a specific foreign particle, such as a surface protein of the parasite T. brucei, the antibody has acquired this novel, ontogenetic function too: "It never lost its generic function, to bind with antigens for the purpose of eliminating them, but it now has a new function super-imposed on the first." (2019, p. 71).

Now suppose that, in this particular case, the antibody wasn't ontogenetically selected for attacking a particular antigen. Instead, the antibody was selected—in error—for attacking a cell or protein within the organisms own body. This is an autoimmune reaction. Antibodies that attack the organisms own tissues are known as 'autoantibodies' and are the dysfunction at the heart of a range of autoimmune diseases. Indeed, research has indicated that pathogenic high-affinity autoantibodies, of the kind that occur in rheumatoid arthritis, systemic lupus and type 1 diabetes "emerge through a process of somatic hypermutation, class-switch DNA recombination and antigen-driven clonal selection." (Elkon and Casali 2008, p. 492). Accordingly, where an antibody has been selected for an autoimmune reaction, the antibody's ontogenetic function (attacking the body's own tissues) is superimposed on the antibody's evolved function of binding to foreign particles for the purposes

[20] This seems to be implied in Garson: he defines antigens as 'foreign particles' and writes that the generic, evolved function of any antibody is to "bind to antigens for the ultimate purpose of eliminating them from the body" (2019, p. 70). However, it is not quite so simple as research show that some antibodies engage in some autoimmune 'housekeeping' which is likely naturally selected (Elkon and Casali 2008).

of eliminating them from the body. In other words, there is a conflict between the antibody's autoimmune ontogenetic function and its evolved function of expelling antigens.

Cancer cells

It has been observed by many theorists that populations of cancer cells display natural selection-like dynamics wherein those cells that have features which enable rapid replication are selected via differential reproduction. Garson notes this phenomenon as well: "A population of cancer cells can exhibit a diverse array of 'adaptive strategies' to subvert the normal barriers to unregulated somatic cell multiplication." (2017, p. 1100). In such a case, the strategies employed by the cancer cell to enable uncontrolled replication would count as novel ontogenetic functions on Garson's view.

However, cancer cells—in virtue of being mammalian cells—also have evolved functions. Due to the constant threat of DNA lesions (caused by various intrinsic and extrinsic stressors, such as cigarette smoke or UV radiation), cells have evolved strategies to respond to damage (Lord and Ashworth 2012). Collectively, these strategies are referred to as the DNA damage response. If the damage to the DNA is relatively minor, the cell may be able to repair the damage and continue to function as a normal cell. However, in the presence of extensive irreparable damage, cells also have the nuclear option—suicide.

> In addition to being mortal, most animal cells can also be suicidal, meaning that they bear mechanisms whose physiological role is to cause their own death. One such physiological cell suicide process is termed apoptosis or programmed cell death.
>
> (Gerl and Vaux 2005, p. 263)

That is, apoptosis or programmed cell death is among the biological functions of normal mammalian cells. However, in some cancers, this process is disrupted. Fernald and Kurokawa suggest that cancer cells likely employ multiple strategies for evading apoptosis: "In normal cells, genotoxic and cytotoxic stress can induce expression of proapoptotic genes … In cancer cells, this mechanism is often nullified by mutation and silencing of key apoptotic genes." (Fernald and Kurokawa 2013, p. 625).

Accordingly, when a cancer cell is failing to self-destruct as it was designed by genetic selection to do, the following obtains: some particular cell (trait T) is failing to die a programmed cell death (effect E) and E is a function (ontogenetically speaking) and a dysfunction (evolutionarily speaking). In other words, there is a conflict between the cells' evolved function and its novel ontogenetic function.

Feral children

Finally, I shall consider one possible cluster of cases. The evidence is less clear in these cases than in those discussed in the above. However, I think they indicate an

important point of departure between Garson and myself; I take it that development can go awry even if the mechanisms of selection are intact. It seems to me that Garson, at best, rather obscures or neglects to engage with this possibility. Development 'gone wrong' is most evident in the most extreme of cases; so-called feral children.

In the most basic sense, feral children are individuals whose brains and bodies have developed, particularly during crucial stages, in such a deprived, artificial or abnormal environment that they fail to fully develop important brain functions, such as language, social skills and motor control. One of the best documented such cases is that of Genie—a pseudonym given to woman discovered at thirteen in 1970s California. Genie had been subjected to extreme abuse, neglect and almost complete social deprivation. Most of her life had been spent tied down in a dark room where she was prohibited from any vocalisation, moving around or interacting with her mother and brother.

Genie displayed many neurodevelopmental problems. Despite attempts at rehabilitation, she never acquired language, and struggled with many aspects of motor control, including chewing, swallowing and walking. She also displayed striking abnormalities in left/right ear language recognition indicating right-hemisphere lateralisation: "The degree of ear advantage resembles that found in split-brains and hemispherectomies, where … only one hemisphere is functioning." (Curtiss 1977, p. 216). In other words, in some respects, Genie's brain was functioning as if she had brain damage.

To what extent were Genie's deficiencies and abnormalities the result of ontogenetic selection mechanisms operating within an extremely unusual and deprived environment? There is no way to know for sure. It is harder still to establish a direct conflict, as I attempt in the above. However, it seems possible that neural selection mechanisms, given am environment during development which departs sufficiently from the 'selective environment' for normal human brain development, can yield impairments which amount to evolutionary dysfunctions—whether or not they were neurally selected.

Reverse conflicts

The above are all cases where trait T is functioning (ontogenetically speaking) but failing to function (evolutionarily speaking). However, conflicts can also occur the other way around.

Because it forms part of the parity of reasoning argument, Garson is keen to maintain that ontogenetically selected effects are appropriately normative—in other words, that they *can* dysfunction. Accordingly, the neural trait that has the ontogenetically selected function of Substance Use Disorder is, per Garson, dysfunctional if it cannot yield this synaptically selected function. Suppose then that someone with Substance Use Disorder (i.e. someone whose brain is yielding Garson's ontogenetic function, but not one or more of its evolutionary functions) recovers from the mental disorder and regains normal, healthy decision-making capacities. If so, then this person (whose brain is functioning precisely as designed by evolution) is still dysfunctional on Garson's view, because it is failing to yield an ontogenetically selected function—substance abuse.

What to make of this? Garson might be tempted to retort that the recovered addict is *not* dysfunctional for there must have been selection for recovery, throughout treatment, such that the original ontogenetic function is now *vestigial* (selected for in the past, but not in recent history) rather than *functional* and, therefore, its failure does not count as a dysfunction. But it need not be so. Recall that—per Garson's own insistence—it is not every neural change which counts as selection and produces new functions. Suppose that recovery came about by cutting-edge precision neural surgery. If so, then Garson would have to concede that the neural trait in question is dysfunctional (in respect of its ontogenetic function) and functional (in respect of its evolutionary function) on his view.

Similar cases of 'reverse conflict' could be imagined in respect of autoantibodies, cancer cells and feral children. Suppose a cancer patient is given a drug which causes the excessively replicating cancer cells to behave like normal cells, halting the growth of the tumour. There is still a conflict of functional norms, for the cancer cells are not yielding their ontogenetic function of uncontrolled replication. Arguably, these cases point to a new problem for Garson. Garson bills the Generalised Theory as sensitive to and concordant with ordinary biological usage.[21] Would any biologist, neurologist or biomedical researcher agree with Garson that treating drug addiction, neutralising pathogenic autoantibodies or halting the progression of a cancer causes biological dysfunction? If not, then the Generalised Theory may fail to live up to the desiderata inherent to Garson's methodology.

Conclusion

I have argued that Garson's critique of the Dysfunction Account from his Generalised Theory of biological function is unsuccessful and reveals an underlying theoretical flaw inherent to the Generalised Theory. Firstly, Garson fails to show that Substance Use Disorder is functional and *therefore* not dysfunctional. At best, he succeeds in showing that Substance Use Disorder is *both* functional *and* dysfunctional. As revealed by the case of Substance Use Disorder, the Generalised Theory yields the conflicting attributions of function and dysfunction. Given the relative fragility of Garson's supporting argument, this is sufficient grounds to reject the Generalised Theory outright.

What remains for Garson? I shall close by suggesting that abandoning the *generalised* part of his theory of evolutionary and ontogenetic functions may bring benefits for Garson's own philosophical objectives. Conflicting functional norms are not the only problem plaguing the Generalised Theory, and it may be that some tricky conceptual stumbling blocks could be circumvented by simply abandoning his

[21] Garson says this about his methodology: "I think the best way to approach functions is to look at ordinary biological usage—that is, how biologists talk about them—as that talk is captured in solve scientific sources. … When biomedical researchers say that a trait is *dysfunctional* they're often indicating, in a pragmatic kind of way, that they trait is … the kind of thing you might want to fix or replace" (Garson 2019, p. 23).

commitment to generality.[22] Doing this could also help facilitate the development of a more thorough, in-depth and specific theory of ontogenetic neural selection in the brain, drawing on Garson's many existing insights and contributions in this area of philosophy and neuroscience. This would no doubt be a valuable contribution to the literature in and of itself, more so, perhaps, than a Generalised Selected Effects Theory of biological function.

Declarations

Conflict of interest This research was conducted while in a receipt of a doctoral studentship funded by the London Arts and Humanities Partnership. The author has no competing interests to declare that are relevant to the content of this article.

References

Bechara A, Van Der Linden M (2005) Decision-making and impulse control after frontal lobe injuries. Curr opin neurol 18(6):734–739

Bolla KI, Eldreth DA, London ED, Kiehl KA, Mouratidis M, Contoreggi CEEA et al (2003) Orbitofrontal cortex dysfunction in abstinent cocaine abusers performing a decision-making task. Neuroimage 19(3):1085–1094

Boorse C (1976) Wright on functions. Philos Rev 85(1):70–86

Boorse C (1977) Health as a theoretical concept. Philos Sci 44(4):542–573

Cooper IR (2020) The concept of disorder revisited: robustly value-laden despite change: Aristotelian Society Supplementary Volume, 94(1):141–161, Oxford University Press

Cummins R (1975) Functional analysis. J Philos 72(20):741–765. https://doi.org/10.2307/2024640

Curtiss S (1977) A psycholinguistic study of a modern-day "wild child". Academic Press, Cambridge

Dennenmoser S, Christy JH (2013) The design of a beautiful weapon: compensation for opposing sexual selection on a trait with two functions. Evolution 67(4):1181–1188

Ekhtiari H, Victor TA, Paulus MP (2017) Aberrant decision-making and drug addiction—how strong is the evidence? Curr Opin Behav Sci 13:25–33

Elkon K, Casali P (2008) Nature and functions of autoantibodies. Nat Clin Pract Rheumatol 4(9):491–498

Fernald K, Kurokawa M (2013) Evading apoptosis in cancer. Trends Cell Biol 23(12):620–633

Garson J (2014) The biological mind: a philosophical introduction. Routledge, Oxfordshire

[22] Because of generality, Garson has to give an analysis of 'population' which is broad enough to apply both to retained synapses in the brain and reproducing organisms (see Garson 2019). He attempts this by requiring a high degree of fitness relevant interactions among members which, concerningly, would seem to make members of distinct species which interact through predator prey relations (in some ecosystem) a paradigmatic biological population. If he gave up on providing a general account, he could simply employ the notion of population which best fit each case.

Garson J (2017) A generalized selected effects theory of function. Philos Sci 84(3): 523–543
Garson J (2018) How to be a function pluralist. Br J Philos Sci 69(4):1101–1122
Garson J (2019) What biological functions are and why they matter. Cambridge University Press, Cambridge
Garson J (2021) The developmental plasticity challenge to wakefield's view. In: Faucher L, Forest D (eds) Defining mental disorder: Jerome wakefield and his critics. MIT Press, Cambridge
Garson J, Papineau D (2019) Teleosemantics, selection and novel contents. Biol Philos 34(3):36
Gerl R, Vaux DL (2005) Apoptosis in the development and treatment of cancer. Carcinogenesis 26(2):263–270
Godfrey-Smith P (1993) Functions: consensus without Unity. Pacific Philoso Q 74:196–208
Godfrey-Smith P (1994) A modern history theory of functions. Noûs 28(3):344–362
Grant S, Contoreggi C, London ED (2000) Drug abusers show impaired performance in a laboratory test of decision making. Neuropsychologia 38(8):1180–1187
Griffiths PE, Matthewson J (2018) Evolution, dysfunction, and disease: a reappraisal. Br J Philos Sci 69(2):301–327
Hogarth L, Balleine BW, Corbit LH, Killcross S (2013) Associative learning mechanisms underpinning the transition from recreational drug use to addiction. Ann N Y Acad Sci 1282(1):12–24
Koffarnus MN, Kaplan BA (2018) Clinical models of decision making in addiction. Pharmacol Biochem Behav 164:71–83
Koob GF, Simon EJ (2009) The neurobiology of addiction: where we have been and where we are going. J Drug Issues 39(1):115–132
Koob GF, Volkow ND (2016) Neurobiology of addiction: a neurocircuitry analysis. Lancet Psychiatry 3(8):760–773
Lemoine M (2013) Defining disease beyond conceptual analysis: an analysis of conceptual analysis in philosophy of medicine. Theor Med Bioeth 34(4):309–325
Lord CJ, Ashworth A (2012) The DNA damage response and cancer therapy. Nature 481(7381):287–294
Lucantonio F, Stalnaker TA, Shaham Y, Niv Y, Schoenbaum G (2012) The impact of orbitofrontal dysfunction on cocaine addiction. Nat Neurosci 15(3):358–366
Matthewson J, Griffiths PE (2017) Biological criteria of disease: Four ways of going wrong. J Med Philos 42(4):447–466
Millikan RG (1989) In defense of proper functions. Philos Sci 56(2):288–302
Millikan RG (1999) Wings, spoons, pills, and quills: a pluralist theory of function. J Philos 96(4):191–206.
Nanay B (2010) A modal theory of function. J Philos 107(8):412–431
Neander KL (1983) Abnormal psychobiology: a Thesis on the 'anti-psychiatry Debate' and the Relationship Between Psychology and Biology (Doctoral dissertation, La Trobe University)
Neander K (1991) Functions as selected effects: the conceptual analyst's defense. Philosophy of Sscience 58(2):168–184
Neander K (1995) Misrepresenting & malfunctioning. Philos Stud 79(2):109–141
Neander K (1998) Mental illness, concept of. Routledge Encyclopaedia of Philosophy, Oxfordshire
Neander K (2017) A mark of the mental: defense of informational teleosemantics, MIT Press
Neander K, Rosenberg A (2012) Solving the circularity problem for functions: a response to Nanay. J Philos 109(10):613–622
Ouellet J, McGirr A, Van den Eynde F, Jollant F, Lepage M, Berlim (2015) MT Enhancing decision-making and cognitive impulse control with transcranial direct current stimulation (tDCS) applied over the orbitofrontal cortex (OFC): a randomized and sham-controlled exploratory study. J Psychiatry Res 69: 27–34
Schoenbaum G, Shaham Y (2008) The role of orbitofrontal cortex in drug addiction: a review of preclinical studies. Biol Psychiatry 63(3):256–262
Schwartz PH (2014) Reframing the disease debate and defending the biostatistical theory. J Med Philos 39(6):572–589
Stalnaker TA, Cooch NK, Schoenbaum G (2015) What the orbitofrontal cortex does not do. Nat Neurosci 18(5):620–627
Torregrossa MM, Quinn JJ, Taylor JR (2008) Impulsivity, compulsivity, and habit: the role of orbitofrontal cortex revisited. Biol Psychiat 63(3):253–255
Wakefield JC (1992) The concept of mental disorder: on the boundary between biological facts and social values. Am Psychol 47(3):373

Wakefield JC (2014) The biostatistical theory versus the harmful dysfunction analysis, part 1: is part-dysfunction a sufficient condition for medical disorder? J Med Philos 39(6):648–682

Zhang XL, Shi J, Zhao LY, Sun LL, Wang J, Wang GB, Epstein DH, Lu L (2011) Effects of stress on decision-making deficits in formerly heroin-dependent patients after different durations of abstinence. Am J Psychiatry 168(6):610–616