Garson, J. (forthcoming). There are no ahistorical theories of function. *Philosophy of Science*

Title: There Are No Ahistorical Theories of Function

Author: Justin Garson

Abstract: Theories of function are conventionally divided up into historical and ahistorical ones. Proponents of ahistorical theories often cite the *ahistoricity* of their accounts as a major virtue. Here, I argue that none of the mainstream "ahistorical" accounts are actually ahistorical. All of them refer, implicitly or explicitly, to history. In Boorse's goal-contribution account, history is latent in the idea of statistical-typicality. In the propensity theory, history is implicit in the idea of a species' natural habitat. In the causal role theory, history is required for making sense of dysfunction. I elaborate some consequences for the functions debate.

Acknowledgments: I'm grateful to audience members at PSA 2018, where I presented this material. I also thank Daniel Dennett and Paul Griffiths for useful feedback.

1. Introduction

Theories of function are conventionally divided up into two main categories, historical and ahistorical (or backwards-looking and forwards-looking). The selected effects theory (Neander 1983, 1991; Millikan 1984) is an example of a *historical* theory, but there are other historical theories, including some versions of the organizational theory (McLaughlin 2001). *Ahistorical* theories include Boorse's goal-contribution account (1976; 1977; 2002), the propensity theory (Bigelow and Pargetter 1987), and the causal role theory (Cummins 1975; Craver 2001; Hardcastle 2002). In the 1970s and 1980s, it was common to see these two sorts of theories as competing with each other, though more recently, philosophers of biology have generally adopted a pluralistic stance, and see them as capturing different aspects of ordinary biological usage (Garson 2018). Still, the validity of the basic distinction has never been seriously challenged.

Many proponents of ahistorical theories have argued that we should accept their theories precisely *on account of* their being ahistorical. In other words, their alleged ahistoricity is often touted as a significant selling point of their theories, and a strong reason to prefer them over historical ones. There are two arguments along these lines. The first argument appeals to bald intuition, and says it's just obvious that functions don't always need history. One fanciful variant of this argument appeals to science fiction cases, like swamp creatures, instant lions, and randomly-generated worlds (e.g, Boorse 1976, 74; Bigelow and Pargetter 1987, 188). But one doesn't have to go as far as science fiction to find plausible cases of ahistorical functions in biology. Many philosophers have a strong intuition that, the very first time a new biological trait emerges and begins to benefit the organism, it has a *function* even if it was never selected for (e.g., Boorse 2002, 66; Bigelow and Pargetter 1987, 195; Walsh and Ariew 1996, 498). The second argument, which is closely related, appeals to ordinary biological usage instead of intuition. It says that historical theories run against the way biologists ordinarily think and talk about functions. At least sometimes, when biologists attribute functions to traits, they neither *cite* nor *refer to* nor *think about* history or evolution (e.g., Godfrey-Smith 1993, 200; Amundson and Lauder 1994, 451; Walsh 1996, 558; Boorse 2002, 73). Hence, ahistorical theories capture important strands of real biology.

In light of the above, my thesis might come as a bit of a shock. I claim that *there are no ahistorical theories of function* – or, put more precisely, the mainstream versions of the allegedly ahistorical theories on the market aren't actually ahistorical. If we poke and prod at those theories a bit, a historical element falls out, like contraband stashed away in a suitcase. In Boorse's version of the goal-contribution account, history is explicitly embedded in his notion of a *statistically-typical* contribution to fitness. In the propensity account, history is embedded, a little less explicitly, in the idea of a species' *natural habitat*. Finally, the only way the causal-role theorist can hope to make sense of dysfunction is to appeal to history.

Before I move on, there is one big qualification I must get out of the way. One could *invent* a purely ahistorical theory of function. One could assert, for example, that *all* of a trait's effects are its functions. In fact, the biologists Bock and von Wahlert (1965, 274)

proposed a theory of function very much along these lines. This theory (pan-functionalism?) would be ahistorical, to be sure, since even if the world were created two seconds ago in pretty much its present form, things would still have effects, and so they'd still have functions. In fact, sometimes scientists actually *do* use the word "function" synonymously with "effect." They say things like, "climate change is a *function* of deforestation," or "poor academic performance is a *function* of malnutrition." But this isn't the ordinary biological use, which the theories I cite above are trying to capture. I'll come back to this point in the conclusion.

So, I need to amend my thesis slightly. Instead of saying that there are no ahistorical theories of function, I want to say that any theory of function that satisfies two very minimal, very traditional, and largely uncontroversial, adequacy conditions, is *also* a historical theory. First, the theory should capture some distinction between functions and accidents (the function of the nose is to help us breathe but not hold up glasses). Second, the theory should capture the possibility of malfunctioning or dysfunction. If my heart seizes up due to cardiac arrest, it's failing to perform its function or it's dysfunctional. All of the theorists I engage with in this paper purport to satisfy these two adequacy criteria, or something in their vicinity, so I'm not begging any questions by insisting on these conditions.

Here's the plan for the rest of the paper. The next three sections will examine Boorse's goal-contribution theory, the propensity theory, and the causal-role theory, in turn. In the conclusion, I'll draw out the big consequences for thinking about functions.

2. Boorse's Goal-Contribution Account

Boorse's view (1976; 1977; 2002), at the most general level, is a goal-contribution account. It holds that a trait's function is just its contribution to a goal. Here, I'll focus on the subclass of functions he calls *physiological* functions. For Boorse, the *physiological* function of a trait is its species-typical contribution to the survival and reproductive prospects of an organism (1977, 555; 2002, 72). (To be more precise, Boorse carves up species into subgroups based on age and sex; the function of a trait is its typical contribution to fitness within the members of that subgroup.) Though he doesn't define a corresponding notion of *dysfunction*, he defines a closely related notion of *disease*: a disease is simply a state that "reduces one or more functional abilities below typical efficiency (1977, 555)."

Neander (1991, 182) raised a now-famous objection against Boorse; she pointed out that Boorse's view, as it stands, can't make sense of pandemic dysfunction: "dysfunction can become widespread within a population…A statistical definition of biological norms implies that when a trait standardly fails to perform its function, its function ceases to be its function; so that if enough of us are stricken with disease (roughly, are dysfunctional) we cease to be diseased, which is nonsense." Pandemic dysfunctions, moreover, don't just occupy the realm of science fiction, as in P. D. James' *The Children of Men*. UV radiation poisoning in anurans is a good example of pandemic dysfunction. Sadly, climate change might create many more pandemic dysfunctions very soon. A good theory

of function shouldn't close off the possibility that all, or most, tokens of a certain trait in a certain species are dysfunctional (or as Boorse prefers, "diseased").

Intriguingly, Boorse doesn't deny the possibility of pandemic disease. Instead, he says that in order to make sense of pandemic disease, one has to appreciate function's *historical depth*. Specifically, when we consider what's "statistically typical" for a trait, we cannot just look at what is typical right now. We have to examine the trait's behavior over a slice of time that includes the present moment and reaches far back into the past: "*Obviously*, some of the species' history must be included in what is species-typical (2002, 99; my emphasis)." He tells us that this time-slice should be longer than "a lifetime or two," and might include "millennia."

This is an extraordinary admission, given that much of Boorse's core argument for his view was propped up on the claim that both biology and intuition need purely ahistorical functions, uncluttered by history. His admission implies that his two key arguments for the view don't work. First, by his own lights, it's not the case that biologists don't refer to history; implicitly, when they talk about what's statistically-typical, they *are* talking about history. Second, regardless of whether or not intuition supports ahistorical functions, Boorse's theory doesn't. It's just not true, on Boorse's account, that if lions popped into being from an unparalleled saltation, their distinctive parts and processes would have functions. They wouldn't, since they don't have the right history (or to be more precise, they have no history at all).

3. The Propensity Theory

Bigelow and Pargetter (1987) also developed an influential "ahistorical" theory of function, the propensity theory. They reject the selected effects theory (and etiological accounts more generally) because the selected effects theory gets the *modality* of functions wrong. In other words, the statement, "functions are selected effects," if true, is contingently true; it might be true on the actual world, but there are possible worlds at which it's false. To illustrate the point, they ask us to consider a world that is pretty much the same as ours except that it randomly popped into being five minutes ago. On that world, they claim, there would still be functions, just no selected effects (188): "we have the intuition that the concept of biological function…[is] not thus contingent upon the acceptance of the theory of evolution by natural selection." This consideration prompts the need for an ahistorical theory.

For Bigelow and Pargetter, functions are propensities, or probabilistic dispositions. We might quibble over what exactly dispositions are, but any good definition will cite three parts: structure, environment, and behavior. Consider the solubility of salt. There is a *structure*, namely, the polar molecular structure composed of sodium and chloride; there is an *environment*, namely, water; there is a *behavior*, namely, dissolving. When we say that salt is disposed to dissolve in water, we're saying that, if you were to take something with this structure, and put it in this environment, it would perform this behavior, all things equal.

Functions, too, are dispositions. Consider "the function of the heart is to circulate blood." For this statement to be true, there must be a structure (the heart, embedded the right way in the circulatory system), an environment (which they call the creature's *natural habitat*), and a behavior (conferring a fitness boost on the organism). If one were to put the structure in its natural habitat, it would increase the fitness of the organism (relative, I suppose, to creatures without hearts). The crucial distinction between their view and Boorse's is that in their view, a trait's function doesn't depend on actual frequencies of performance. A trait needn't have an actual track record of boosting fitness to have a function; a mere propensity will do.

This raises the thorny question of what a creature's *natural habitat* is. For they're clear that a creature's natural habitat isn't just any environment the creature happens to find itself in. Curiously, they refuse to define this crucial notion; instead, they brush it off as vague, but unproblematically so: "there may be room for disagreement about what counts as a creature's 'natural habitat;' but this sort of variable parameter is a common feature of many useful scientific concepts" (192). But one could at least form the suspicion that if one analyzed this unproblematically vague notion, one would find some reference to history tucked away inside of it.

This suspicion is confirmed in the very next paragraph of their paper. There, they tell us that, if a creature's environment were to change very suddenly, then "natural habitat" will still refer to the *old* environment, and not the *new* one (ibid). There's a time lag built into the very idea of a natural habitat. So, for example, if climate change melts enough Arctic ice, then, at least for a time, the polar bear's natural habitat (and by extension, the natural habitat of the trait itself, namely, their thick, water-repellant fur) is the icy habitat of yore and not the contemporary, denuded one. They take that as given, and I agree.

But why would this be? What *makes it the case* that, in cases of rapid habitat change, "natural habitat," at least for a time, refers to the old environment and not the new one? What makes it true, I suspect, is that the idea of a natural habitat is an intrinsically historical notion. It's something like *the environment within which the species recently survived and thrived*. And if that's not what a natural habitat is, I would like to know what it is *such that*, if a creature's actual habitat shifts suddenly, the natural habitat, for a little while, is still the old one. Just because a concept is vague around the edges, that doesn't excuse one from the obligation to give some sort of analysis. Perhaps one could revise the theory and drop all reference to "natural habitat," as suggested by Griffiths (2009, 27), but that remains to be worked out in a rigorous way. Moreover, it's not clear whether such a theory, when rigorously developed, would hang together with the two adequacy criteria.

Hence, I conclude that, contrary to rumor, the propensity theory is not an ahistorical theory, or not demonstrably so. But if that's right, proponents of the propensity theory lose one of the main virtues of the view, which is to get the modality of functions right. To be fair, there's still a sense in which their view *is* ahistorical. What they can do, that the selected effects theorist can't, is to attribute functions to novel traits – so long as that

novel trait belongs to the members of a species that has been around long enough to have a natural habitat. Suppose a gene mutation confers a benefit on an organism, say, pesticide resistance in a flour beetle. I suppose they can say that, at the very moment at which it first confers that benefit, the gene mutation has a function, namely, to make the beetle withstand a certain pesticide. This result, they claim, is "intuitively comfortable" (195). But they can say that only because flour beetles themselves have a history, and so we can talk meaningfully about their natural habitats. Moreover, I think they'll still have a rough time explaining dysfunction (Neander 1991, 183), for reasons I'll point to in the next section. Finally, I think there are good theory-neutral reasons for saying that beneficial traits, on their very first appearance, don't have functions, but rather, whatever benefit they bring is a lucky accident. But I won't argue for that here (see Garson 2019, Chapter 2).

4. The Causal Role Theory

What about the causal role theory of function? This appears to be a purely ahistorical view. The causal role theory says, roughly, that the function of a *component* of a system consists in its contribution, in tandem with the other components, to a system-level capacity of interest (Cummins 1975; Craver 2001; Hardcastle 2002). Craver (2001) helpfully elaborates this view by specifying that the part in question must be a component of a *mechanism*. All of the basic ingredients of this theory, it seems, are ahistorical: capacities, components, organization, hierarchy, interests. Even if the world were created five minutes ago, in pretty much its present form, things would still have causal role functions.

The problem enters when we think about dysfunction. Cummins (1975, 758) insisted that functions are dispositions, or capacities: "…to attribute a function to something is, in part, to attribute a disposition to it." The function of a trait *token*, then, consists in its capacity to contribute to a system-level effect. But what if the token in question, through defect or disease, loses the capacity, and so can't contribute to the system-level effect? Then, by Cummins' analysis, it doesn't have the relevant function – so it can't be dysfunctional either.

Causal role theorists have, by and large, been silent about how to make sense of dysfunctions from this perspective. Almost everything they've had to say on that score, however, is consistent with the following theme: a trait *token* is dysfunctional when it can't do what other trait tokens generally, or typically, do to contribute to the system-level effect of interest. Consider Godfrey-Smith (1993, 200): "Although it is not always appreciated, the distinction between function and *mal*function can be made within Cummins' framework…If a token of a component of a system is not able to do whatever it is that other tokens do, that plays a distinguished role in the explanation of the capacities of the broader system, then that token component is malfunctional." Craver (2001, 72), offers the same general line: "…the ascription of a function to a malformed or broken part is derivative upon a description of how that *type* of part (X) fits into a *type* of higher-level mechanism (S). The malformed and broken part can be identified as an X by

the typical properties and activities of Xs....” This is, at root, to rely on a statistical norm for making sense of dysfunction.

This account of dysfunction, like Boorse’s, stumbles when it encounters the problem of pandemic dysfunction. For the modification suggested above implies that, if everyone’s heart seized up at once, nobody’s heart would have a function anymore, so nobody’s heart would be dysfunctional. The best way to solve this problem, and perhaps the only way, is the way Boorse took, namely, to say that the function of a trait is its typical contribution to some system effect, where what’s typical is assessed over a chunk of time that stretches back into the past, for at least “a lifetime or two,” and perhaps “millennia.” But if causal role theorists take that line, they’d have a historical theory.

Craver (2001) and Hardcastle (2002) suggest, all too fleetingly, a different way of thinking about dysfunction, one that depends not on statistics, but on our values, that is, the values and goals of people who make function attributions. Craver (2001, 72) suggests that traits are dysfunctional when they cannot do what people *want* them to do: “the mechanistic role of the broken part only appears against the fixed backdrop of shared assumptions about a type of mechanism within which parts of this type generally (or preferably) make important contributions.” The parenthetical remark alludes to a substantially new doctrine, one that demands our full concentration. It suggests that dysfunction is a mirror of human preferences and goals, of our wishing and wanting. If my heart seizes up, it’s dysfunctional, since it’s not doing *what I want it to do.*

Hardcastle (2002) makes remarks along similar lines. She first says that the function of a trait – what it’s “supposed to do,” as she puts it – depends on the goals of the scientific discipline that makes the investigation: “The teleological goal for some trait…depends upon the discipline generating the inquiry” (153). The palmomental reflex causes a chin twitch when you stroke an infant’s palm; it’s just an accident of cortical wiring with no deep evolutionary rationale. Still, she says, it has the *function* of indicating the state of brain development in infants, because that’s how biomedical researchers use it. She then says that something is malfunctional just when it cannot do what it’s supposed to do (152). The palmomental reflex is malfunctional when it can’t indicate the state of brain development. Simply put, dysfunction happens when a trait can’t do what we want.

But dysfunctions can’t be reduced to mere preferences in any straightforward way; this is a point that’s been taken in the literature for decades (e.g., Boorse 1977, 544; Wakefield 1992, 372), for reasons that scarcely need to be rehearsed. I’d prefer not to need sleep and water; I’d prefer if nobody had to go through the pain of childbirth or teething, either. But none of those things are dysfunctions. For that matter, I’d prefer if my hands were equipped with retractable adamantium claws. The fact that my hands can’t do what I want them to do doesn’t make them dysfunctional. If one really wanted to run with this value-centered line about dysfunction, one would *at least* have to add that, in order for a trait to be dysfunctional, it’s not enough that it doesn’t do what I prefer, but I must also have a *reasonable expectation* that it *should* act in the way that I prefer. But what could possibly ground a *reasonable expectation* that my hand (say) work in a certain way? Only this: that hands usually *do* work in the preferred way. But then we’re back to statistical

norms, and long historical slices of time. This value analysis of dysfunction isn't a contender to a statistical analysis; instead, the former presupposes the latter.

I've walked through three allegedly ahistorical theories of function, and shown that none of them are purely ahistorical; they're *infected* with history. The conclusion will say what we should do next.

5. Conclusion

There are no ahistorical theories of function, at least among the mainstream theories that are put forward as ahistorical. The first, Boorse's goal-contribution theory, explicitly refers to what's statistically typical for a trait, where what's typical is assessed over a long historical period of time. The second, the propensity theory, refers to the creature's natural habitat, which is implicitly historical. And the third, the causal role theory, can't hope to make sense of dysfunction (or so I argue) without appealing to a statistical norm, and thereby (following Boorse) to history. None of these theories will give functions to the parts of swamp creatures, instant lions, or anything on worlds that are similar to ours except for being randomly generated five minutes ago. The propensity theory, at least, can give functions to novel traits as soon as those traits begin benefiting their bearers, as long as the population in which the traits emerge has been around for long enough to have something like a natural habitat. But even that theory will probably encounter problems when it comes to making sense of dysfunction, though I haven't pushed that line in any detail here.

If my thesis is correct – that there are no ahistorical theories of function – three consequences immediately follow. First, we need to jettison this whole way of dividing up theories of function. The distinction between etiological and non-etiological theories serves us much better. An *etiological* theory holds that function ascriptions either are, or purport to be, causal explanations for the existence of traits. Non-etiological theories hold that function ascriptions are not, and they don't purport to be, causal explanations for traits. But the crucial point is that being etiological and being non-etiological are just *two different ways of being historical*.

Second, given that there are no ahistorical views, the two main arguments that have repeatedly been put forward for those theories – the argument from intuition and the argument from ordinary biological usage – don't actually work. If we took those arguments seriously, they'd count as evidence *against* these allegedly ahistorical theories. That doesn't mean those theories are wrong. It does mean, however, that we need to rethink, from the ground up, the motivation for accepting those theories.

A third consequence is that one popular way of thinking about function pluralism must fail. This sort of pluralist wishes to sort all biological usage under two main umbrella theories, the selected effects theory and the causal role theory. An argument for this sort of pluralism is that it mirrors the two main uses of "function" in biology, the historical sense and the ahistorical sense. If I'm right, this incarnation of the pluralist project can't work either.

True, there are some theories of function I haven't addressed here, which fall a bit outside of the mainstream. Might those come to our rescue? In particular, one might wonder how the *modal theory* of function (Nanay 2010) fares with respect to my analysis. The modal theory holds that a function of a trait *token* depends on that token's behavior on nearby possible worlds, where what's "nearby" depends on our explanatory interests. I agree that this is an ahistorical theory through and through, since what function a trait has, and whether or not it's dysfunctional, depend on what's going on at other possible worlds, rather than the actual past. But it also yields a deeply implausible construal of dysfunction. As Neander and Rosenberg (2012) point out, if the modal theory is right, then many traits that biologists don't think of as dysfunctional, like the trait of lactose-intolerance in most Pacific Islanders, would actually be dysfunctional. So, while the modal theory doesn't violate the *letter* of my second adequacy condition – namely, that it should allow for the possibility of dysfunction – it violates the *spirit* of that condition by carving up functions and dysfunctions in a wildly revisionary way.

Nanay (2012) argues that the fact that function ascriptions are relative to our explanatory interests can somehow lessen the sting of this counterintuitive consequence, but I don't see how that helps. To illustrate the problem, consider Temitope, an evolutionary geneticist who's interested in how human beings might evolve in the near future. Temitope considers a possible world to be "nearby" if, at that possible world, she has a counterpart, and her counterpart's genome differs from hers by only a single point mutation, but the rest of the world is largely the same (yielding at least 3 billion nearby worlds). She reasons that, on some of those possible worlds, some of her traits would do things that enhance her inclusive fitness. For example, we might suppose that there is a possible world at which her body's ability to dissolve arterial plaque is substantially enhanced, one at which she has tetrachromatic vision, and one at which she's resistant to malaria. She realizes, with dismay, that her body's actual ability to dissolve arterial plaque represents a dysfunction. In fact, she realizes that, *relative to her explanatory interests*, she has many more dysfunctions than she ever thought possible. So even if we agree that function ascriptions are tethered to explanatory interests, we still get deeply revisionary consequences. In my reckoning, a theory that hangs together pretty well with ordinary biological usage is better than a deeply revisionary one, all things equal (see Garson 2016, 105-7, for further discussion).

There's a twist to my story, which I alluded to in the introduction. I think there is a prominent sense of "function" in scientific circles that is ahistorical. Consider that climate change is a function of deforestation, poor academic performance is a function of malnutrition, and wildlife habitat is a function of soil. These notions are ahistorical through and through. "Function," in this context, means little more than "effect," and perhaps (as in the last of the three examples) "helpful effect." But this tepid sense of function isn't going to sustain a distinction between function and accident, nor will it give us any sense of dysfunction. This is the sort of "function" that Bock and von Wahlert (1965, 274) were getting at when they equated functions with "all physical and chemical properties arising from [the trait's] form." It's also the sort of "function" that Neander (2017) describes in her recent discussion of "minimal functions." But the proponents of

the allegedly ahistorical theories want functions to do much more than that. They are trying to capture the ordinary biological sense (or *an* ordinary biological sense) of "function," where functions differ from accidents and sometimes things are dysfunctional. Unfortunately, they can't have what they want.

References

Amundson, R., and G. V. Lauder. 1994. Function without purpose: The uses of causal role function in evolutionary biology. *Biology and Philosophy* 9: 443-469.

Bigelow, J., and Pargetter, R. 1987. Functions. *Journal of Philosophy* 84: 181-196.

Bock, W. J., and von Wahlert, G. 1965. Adaptation and the form-function complex. *Evolution* 19: 269-299.

Boorse, C. 1976. Wright on functions. *Philosophical Review* 85: 70-86.

Boorse, C. 1977. Health as a theoretical concept. *Philosophy of Science* 44: 542- 573.

Boorse, C. 2002. A rebuttal on functions. In *Functions: New Essays in the Philosophy of Psychology and Biology*, ed. A. Ariew, R. Cummins, and M. Perlman, 63-112. Oxford: Oxford University Press.

Craver, C. 2001. Role functions, mechanisms, and hierarchy. *Philosophy of Science* 68: 53–74.

Cummins, R. 1975. Functional analysis. *Journal of Philosophy* 72: 741–765.

Garson, J. 2016. *A Critical Overview of Biological Functions*. Dordrecht: Springer.

Garson, J. 2018. How to be a function pluralist. *British Journal for the Philosophy of Science* 69: 1101-1122.

Garson, J. 2019. *What Biological Functions Are and Why They Matter*. Cambridge: Cambridge University Press.

Godfrey-Smith, P. 1993. Functions: Consensus without unity. *Pacific Philosophical Quarterly* 74: 196-208.

Griffiths, P. 2009. In what sense does 'nothing make sense except in the light of evolution'? *Acta Biotheoretica* 57: 11-32.

Hardcastle, V.G. 2002. On the normativity of functions. In *Functions: New Essays in the Philosophy of Psychology and Biology*, ed. A. Ariew, R. Cummins, and M Perlman, 144-156. Oxford: Oxford University Press.

Millikan, R. G. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.

Nanay, B. 2010. A modal theory of function. *Journal of Philosophy* 107: 412-431.

Nanay, B. 2012. Function attribution depends on the explanatory context: A reply to Neander and Rosenberg's reply to Nanay. *Journal of Philosophy* 109: 623-627.

Neander, K. 1983. *Abnormal Psychobiology*. Dissertation, La Trobe.

Neander, K. 1991. Functions as selected effects: The conceptual analyst's defense. *Philosophy of Science* 58: 168–184.

Neander, K. 2017. Functional analysis and the species design. *Synthese* 194: 1147-1168.

Neander, K., and Rosenberg, A. 2012. Solving the circularity problem for functions. *Journal of Philosophy* 109: 613-22.

Wakefield, J. C. 1992. The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist* 47: 373–388.

Walsh, D.M. 1996. Fitness and function. *British Journal for the Philosophy of Science* 47: 553-574.

Walsh, D. M., and A. Ariew. 1996. A taxonomy of functions. *Canadian Journal of Philosophy* 26: 493-514.

Wright, L. 1973. Functions. *Philosophical Review* 82: 139-168.