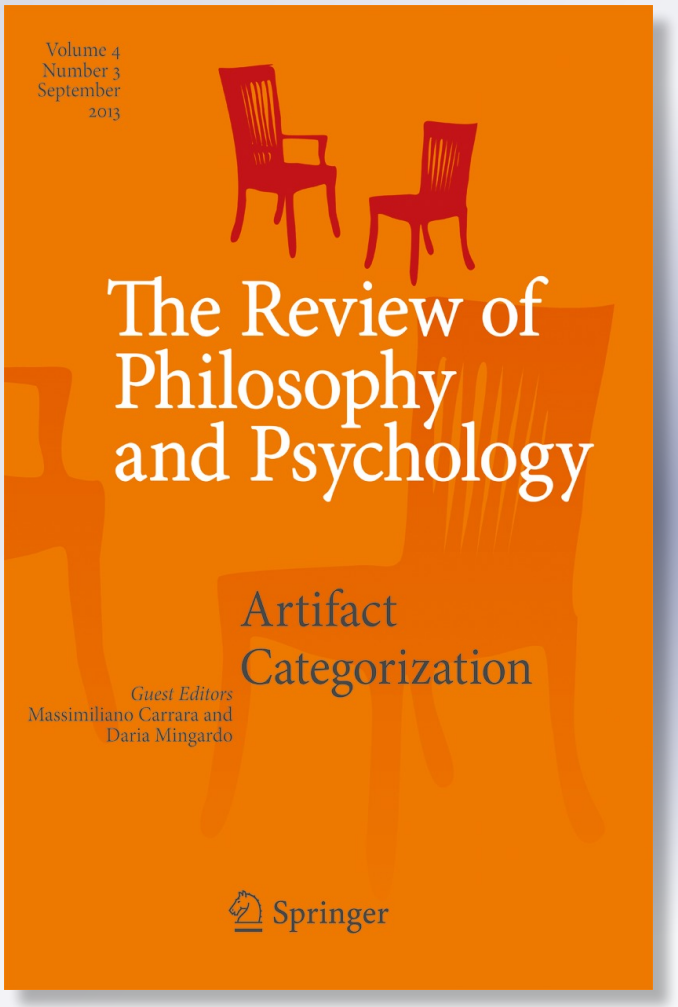
Artifact Categorization and the Modal Theory of Artifact Function

Bence Nanay

Review of Philosophy and Psychology

ISSN 1878-5158
Volume 4
Number 3

Rev.Phil.Psych. (2013) 4:515-526
DOI 10.1007/s13164-013-0143-6



Springer

Rev.Phil.Psych. (2013) 4:515–526
DOI 10.1007/s13164-013-0143-6

Artifact Categorization and the Modal Theory of Artifact Function

Bence Nanay

Published online: 12 May 2013

Abstract Philosophers and psychologists widely hold that artifact categories – just like biological categories – are individuated by their function. But recent empirical findings in psychology question this assumption. My proposal is to suggest a way of squaring these findings with the central role function should play in individuating artifact categories. But in order to do so, we need to give up on the standard account of artifact function, according to which function is fixed by design, and replace it with a more context-sensitive account, according to which function attributions depend on the explanatory context and, as a result, categorization also depends on the explanatory context, just as the empirical findings from psychology suggest. I argue that the recent 'modal theory of function' originally proposed to explain biological function, is capable of providing such an account of artifact function.

1 Introduction

How do we individuate artifacts? What makes, say, corkscrews different from non-corkscrews? One influential view both among philosophers and among psychologists is that artifacts are individuated by means of their function – what makes a corkscrew a corkscrew is that it has a certain function – it has the function to open wine bottles (for the most sophisticated versions of this claim, see Rips 1989; Ahn 1998; Ahn et al. 2001; Bloom 1996, 1998, 2007).

But there is very convincing recent empirical evidence that we in fact do not individuate artifacts by means of their function. At least we don't always do so: these new empirical results suggest a more pragmatist picture: the way we categorize artifacts depends very much on what we are interested in – sometimes it is their function, sometimes it is something else. Hence, we do not individuate artifacts only according to their function (see Malt and Johnson 1992; Malt 1994, 1995; Sloman and Malt 2003, 2007; Chaigneau et al. 2004).

How can we resolve this disagreement? One obvious way to proceed would be to just reject the old philosophical assumption that function plays an essential role in

B. Nanay (✉)
University of Antwerp, Antwerp, Belgium and Peterhouse, Cambridge University, Cambridge, UK
e-mail: bence.nanay@ua.ac.be

individuating artifacts. But I want to proceed differently. I will argue that some of this disagreement on artifact categorization is relatively easy to dissolve – they only disagree if we assume certain accounts of function, ones that we have good reasons to reject anyway.

More precisely, both the proponents and the opponents of the idea that artifacts are individuated by means of their function assume that the function of artifacts is fixed by design. Philosophers are very explicit about this assumption. As Philip Kitcher put it, "the function of S is what S is designed to do". (Kitcher 1993, p. 380, see also Millikan 1984, p. 17 and Williams 1966, p. 209).[1] But psychologists also make this assumption albeit sometimes implicitly (a strikingly explicit treatment is Sloman and Malt 2003, esp. p. 565).

My suggestion is to keep the claim that artifacts are individuated by means of their function and reject the extra assumption that function is fixed by design. If we replace this account of function with a different one – one inspired by biological function, more precisely, the 'modal theory of function' (Nanay 2010, 2011, 2012, see also Neander and Rosenberg 2012) – then we can accept, with the philosophers, that function individuates artifact categories and we can also accept, with the psychologists, that artifact categorization is sensitive to context and to our pragmatic interests. The reason for this is that function attribution itself, at least according to the 'modal theory of function', is also sensitive to context and to our pragmatic interests.

2 Biological Function Versus Artifact Function

The function of my corkscrew is to open bottles. The function of my heart is to pump blood. These two function-attributions are of different kind. My corkscrew is an artifact, whereas my heart is a biological trait. Artifact function seems to be the easier of the two kinds to analyze. The standard way of explaining artifact function is with reference to the notion of design. My corkscrew has the function to open wine bottles if and only if it was designed to open wine bottles. If we wonder what the function of an artifact may be, we should just ask the designer to get an answer (Kitcher 1993, p. 380, see also Millikan 1984, p. 17 and Williams 1966, p. 209).

This explanatory scheme will not work in the case of biological functions, as there is no designer who designed biological traits. Thus, it seems that in spite of the fact that we talk about functions both in the artifact and in the biological case, these two kinds of function are very different indeed: one is fixed by design, whereas the other is not.

This asymmetry has been bothering most philosophers who are thinking about function. If my heart and my corkscrew both have function, we must be able to come up with a theory of function that covers both kinds of case. Even if the attribution of biological an artifact function is somewhat different, they must have something in common. But if one of them is determined by design, whereas the other is not, it is not clear what this common denominator could be.

The standard way of resisting this asymmetry is to accept that artifact function is fixed by design and argue that something similar could be said of biological function. Many of the most influential attempts to characterize biological function take an

[1] Even those who aim to reconsider the role the notion of design plays in the explanation of biological function (for example, Allen and Bekoff 1995; Buller 2002) accept a weaker claim that if x is designed to do F, then the function of x is to do F. For dissenting views, see Preston 1998, Vermaas-Houkes 2003, 2004.

implicit account of artifact function-attributions for granted and use considerations about what is assumed about artifact function to arrive at a viable theory of biological function. More precisely, artifact function is fixed by past design, thus, biological function must be determined by something similar, that is, past selection.

I will follow the opposite route. I argue that biological function is not determined by past selection, but by modal facts and then proceed to extend this account to artifact function: artifact function is also determined by modal facts. In short, we can restore the symmetry between artifact and biological function-attributions, without being committed to the claim that artifact function is fixed by design. Although it is widely assumed in the function literature that artifact function is determined by design (see Kitcher 1993 for a summary), there have been some dissenting views (Dipert 1993; Preston 1998, Vermaas-Houkes 2003, 2004). I would like to add to these dissenting voices and outline a new alternative that is based on what I take to be the correct account of biological function: the modal theory.

The plan for the paper is the following. I point out that the attempt to explain biological function on the model of artifact function, the so-called etiological theory, faces some serious objections (Sections 3 and 4). In order to avoid these objections, we need to define biological function in modal terms (Section 5). Finally, I return to the artifact case and give a parallel modal account of artifact function and argue that if we accept a modal theory of artifact function, we can reconcile the philosophical and the psychological outlook on artifact categorization (Sections 6 and 7).

3 The etiological Theory of Biological Function

The most widespread notion of biological function is the following: a trait of an organism has function F if and only if its performing F has contributed to the survival of the ancestors of this organism. This notion of function is usually referred to as 'etiological': what determines the function of a trait is its history. The function of the human heart is to pump blood because the fact that the heart pumped blood contributed to the survival of our ancestors (Millikan 1984; Neander 1991a, b; Griffiths 1993; Godfrey-Smith 1994, see also Wright 1973).

Note that this theory of biological function restores the continuity between the explanation of biological and artifact functions. The function of my corkscrew is to open wine bottles because it has been designed to open wine bottles, whereas the function of my heart is to pump blood because it has been selected for pumping blood. In both cases, function is fixed by the past: past design or past selection.

The etiological theory of function gives an elegant way of handling the possibility of malfunctioning. An important feature of the concept of function that any theory needs to be able to account for is the possibility of malfunctioning. An object may have a function but fail to perform this function. If my heart skips a beat, it still has the function to pump blood, but at that moment it fails to perform this function: it malfunctions. The same is true of broken corkscrews. If we accept the etiological theory of function, malfunctioning can be accounted for very easily: it is perfectly possible that a trait of an organism has been selected for doing F, but at the moment it does not perform (or maybe couldn't even perform) F.

At present, the etiological theory is the dominant view of function. This, however, does not mean that this theory is generally thought of as unproblematic. Probably the

most famous objection to the etiological view is based on the swampman thought experiment. A very direct consequence of the etiological definition of function is that what fixes the function of a trait is its past, not its present. Hence, if an organism that is molecule for molecule identical to me (the swampman) were created by chance, its organs would not have any functions, since it would lack the evolutionary history that would fix the function of these organs (Neander 1996; Millikan 1996).

Without going into the Byzantine details of the swampman literature (Dennett 1996; Millikan 1996; Neander 1996; Papineau 1996, 2001, Braddon-Mitchel and Jackson 2002), it can be pointed out that it is probably not a fatal objection to the etiological theory of function.[2]

The etiological theory could be thought of as an attempt to apply considerations from the standard account of artifact function to the biological domain. I will argue in the next section that it is unlikely to succeed. My strategy then will be to outline a modal theory of biological function and then apply this account to the domain of artifacts.

4 A Problem for Accounts of Biological Function – the Individuation of Trait Types

I would like to raise another, more general and perhaps more fundamental objection than the swampman thought experiment that should persuade us to discard the etiological theory (see also Nanay 2006, 2010).

The etiological definition of function presupposes that *trait types* can be individuated in an unproblematic manner. The trait whose function is to be defined and the traits that have been selected for in the past must be *of the same type*. But how can we individuate trait types? What makes hearts different from non-hearts?

The problem is that there is no coherent non-circular way of individuating trait types that is available to the etiological theory of function.

According to the etiological theory, the function of my token heart is determined by some facts about what happened to some other traits that were tokens of the same type as my heart: whether they were selected for, and if so, what they were selected for. Thus, the function of my token heart is determined by something that happened to some other tokens of the same type.

But then the etiological definition of function presupposes an independent explanation for how trait types are individuated. The real problem is that no such independent explanation is available. The most plausible candidate for how to individuate trait types uses (at least partly) functional criteria: a token object belongs to trait type T if and only if it has certain functional properties: if it has the function to do F (the most important alternatives are discussed in Nanay 2006, 2010). Those entities are hearts that have the function of pumping blood. Those entities that do not have this function are not hearts. As Karen Neander puts it: "Most biological categories are only definable in functional terms" (Neander 1991a, 180, See also Beckner 1959, 112, Lewens 2004, 99, Burge 1989, 312). But the etiological theory of function cannot help itself to this way of individuating trait

[2] There are two main strategies available for the advocate of the etiological view. The first is to bite the bullet and agree that the heart and kidney of the swampman do not have any function (Neander 1996; Papineau 1996; Millikan 1996). The second is to deny the possibility (or the biological/philosophical relevance) of the emergence of swampman (Dennett 1996; Millikan 1996; Neander 1991b; Papineau 2001).

types when defining function without running into circularity. As we have seen, the etiological definition of function presupposes an account of trait type individuation. Now, if we want to avoid circularity, we cannot use the notion of function in order to explain trait type individuation. When we are explaining function, the claim that x* (the trait whose function we are explaining) is a token of type X (the traits that have been selected in the past) is part of the *explanans*. Hence, we cannot use the *explanandum* (function) to explain part of the *explanans* (why x* is a token of type X).[3]

5 A modal Theory of Biological Function

To sum up the argument so far, the etiological theory should be disposed of, and we should look for some other theory of function. The problem is that *all* alternatives to the etiological theory of function define the function of a token trait in terms of some properties of the trait type this trait is a token of. Hence, the alternatives of the etiological theory – the propensity theory (Bigelow-Pargetter 1987; Mills and Beatty 1979), the relational theory (Walsh 1996) as well as Cummins's 'minimalist' theory (Cummins 1975, 2002) – rely on an independent account of individuating trait types.[4]

[3] This problem is acknowledged by the defenders of the etiological theory of function. See Neander 2002, 403, see also Griffiths 1993; Davies 2000, 2001.

[4] It may not be entirely clear why the propensity theory of function faces the trait type individuation problem. According to the propensity definition, a trait "has a (biological) function just when it confers a survival enhancing propensity on a creature that possesses it" (Bigelow and Pargetter 1987, p. 192). As Peter Godfrey-Smith pointed out, this definition hides a type/token ambiguity and, more generally, the propensity theory oscillates between talking about the propensity of a trait token (see, for example, Bigelow and Pargetter 1987, p. 192) and taking about the propensity of a trait type (see, for example, Bigelow and Pargetter 1987, pp. 194–195). Godfrey-Smith argues, convincingly, that while both readings face problems, the latter is the more plausible among the two (Godfrey-Smith 1994, p. 360). But regardless of what may be the most charitable interpretation of the propensity account, it is worth considering whether a 'token' version of the propensity account could avoid the trait type individuation objection. My answer, unsurprisingly, is that it couldn't, for the following reason. One of the best-known objection to the propensity account is that the fitness of an organism always depends on the environment as well as the other organisms in that environment. Hence, if the function of a trait is identified with its capacity to contribute to the fitness of the organism, then it needs to be specified in what environment and against what other organisms it will do so (Godfrey Smith 1994, pp. 352–353). Further, whether the propensity of a token trait contributes to the organism's fitness also depends on whether the other traits and organs of this organism cooperate. The propensity view wants to fill this gap by referring to the 'natural habitat' of the organism: a trait has a function if it has a capacity to contribute to the fitness of the organism in the 'natural habitat' of the organism. As Bigelow and Pargetter say, "when we speak of the function of a character, therefore, we mean that the character generates propensities that are survival-enhancing in the creature's natural habitat" (Bigelow and Pargetter 1987, p. 192). What they mean by 'natural habitat' encompasses not only the environment and the other organisms in this environment, but also other organs of the same organism. The objection raised by some of the critics of the propensity view (Neander 1991b, Godfrey Smith 1994; Walsh 1996) is that the only way they can make sense of this notion of 'natural habitat' is to identify it with the past habitat of the organism, which would make the propensity view into a version of the etiological view it aims to replace. But for the present purposes we do not even need to accept this conclusion. What matters for us is that when taking about the inter- and intra-organismic 'natural habitat' of a token organism in the definition of the function of a token trait, the propensity account needs to bring in trait types – the trait types of the other traits of the organism that the propensity of the trait in question depends on. Even if we don't hold that the concept of 'natural habitat' is intrinsically problematic, it is sufficient for us to see that it presupposes an independent account of trait type individuation. In short, the propensity view is not in a better position than the etiological one when it comes to avoiding the trait type individuation objection.

How can we then possibly give a plausible theory of function? The only solution I can see is to define function without any reference to trait types. If we could define function without appealing to trait type individuation, then we could use this definition of function to individuate trait types without running into circularity.

But then the function of a token trait must be determined entirely by the properties of *that very trait token* and not by the properties of other tokens of the trait type this token belongs to. How can we explain malfunctioning in this framework? When a trait malfunctions, it is supposed to do (that is, it has the function to do) F, but it does not do F. My heart malfunctions when it does not pump blood (though it is supposed to/it has the function to do so). If we define the function of a trait token in terms of the properties of that trait token alone, then it is difficult to see how the function can be different from what the trait token actually does. In other words, it is difficult to see how such an account of function could explain malfunctioning.

In the light of these constraints, options are fairly limited with regards to an unproblematic definition of biological function. In fact, I can see only one such option, namely, to attribute modal force to claims about function. To put it simply, trait x may not perform F, but if it were to perform F, this would contribute to the survival of the organism with x. So the basic idea is that the function of x is F if and only if it is true that if x is doing F, then this *would* contribute to the survival of the organism with x. But some clarifications, comments and elaborations are in order.

First, the talk about contribution to the survival of an organism, which is a standard way of analyzing function, is vague. What really matters in natural selection is not the survival, but the inclusive fitness of the organism. Second, I defined function with the help of a counterfactual. Any theory of counterfactuals could be used to fill in the details of this definition, but, for simplicity, we can use Lewis' theory (Lewis 1973): the function of organism O's trait x is to do F at time t if and only if some 'relatively close' possible world (different from the actual world) where x is doing F at t and this contributes to O's inclusive fitness are closer to the actual world than any of those possible worlds where x is doing F at t, but this does not contribute to O's inclusive fitness (see Nanay 2010 for a detailed defense of this definition).[5]

If x is not doing (or even cannot do) F in the actual world, but in a 'relatively close' possible world it is doing F and its doing F contributes to the organism's inclusive fitness, then we can still attribute function F to x. This is exactly what happens if a trait is malfunctioning: if it fails to perform its function.

How can this proposal deal with the cases that are problematic for the etiological approach? If the swampman's heart pumped blood, then this would contribute to the inclusive fitness of the swampman (this follows from the supposition that the swampman

[5] While Lewis's theory is useful in many ways – especially in clarifying how function attribution depends on the explanatory context in terms of what counts as a 'relatively close' possible world, see esp. Nanay 2012 – it is somewhat misleading in some other ways. If x is doing F in the actual world and this contributes to O's inclusive fitness, then this, being the actual world, is closer than the closest possible world where x is doing F but this does not contributes to O's inclusive fitness. Should we then conclude that whatever x is doing in the actual world in such a way that it contributes to O's inclusive fitness automatically makes it x's function? No we shouldn't. This is why the definition put a restriction on the 'relatively close' possible, *but not actual*, worlds. What this means is that x may or may not be doing F in the actual world, but when looking for the closest possible worlds where it's doing F contributes to O's inclusive fitness, we should ignore the actual world. This point will play an important role in Section 7(a). Thanks for an anonymous referee for pushing me on this issue.

is molecule by molecule identical to a human being), hence, the swampman's heart has the function to pump blood, in spite of the fact that he lacks history.

Finally, the modal theory of function is obviously not vulnerable to the trait-type individuation objection, because it does not use trait types when defining function. It defines the function of a token trait entirely in terms of the properties of this token trait. To sum up, if we conceive of function the way I suggested, some of the worrying consequences of the etiological view disappear.

6 A Modal Theory of Artifact Function

So far, I outlined an account of biological function. However, if this theory is correct, then the explanation of biological function is very different from the standard explanation of artifact functions. Artifact function is widely held to be fixed by design, whereas biological function is fixed by modal facts.

One of the attractions of the etiological theory of function was that it can provide a theory of biological function that is continuous with the way we usually explain artifact function (see Kitcher 1993). I will follow the opposite route. Instead of constructing a theory of biological function that would mirror the standard way of thinking about artifact function, we should reevaluate the standard understanding of artifact function. In short, my claim is that artifact function is not, or at least not always, fixed by design. Let me begin with an example.

The slinky was not designed to be used as a toy that can 'walk' downstairs. It was designed, apparently, to be a tension spring in a horsepower monitor for naval battleships. Nonetheless, its function now is to 'walk' downstairs. Other examples include truck tires used for football practice and old chalkboards used as dinner tables in some trendy households.

Thus, it is not true of artifacts in general that x has function F if and only if x was designed to do F. But then how can we explain artifact function?

My suggestion, not surprisingly, is that function attribution to artifacts depends on modal facts about the token artifact. The function of an artifact is fixed by what *would* contribute to the fulfillment of the goals of the agent who is using the artifact.[6] This could be spelled out in the following way: artifact x has function to do F at time t if and only if some 'relatively close' possible (but not actual) world where x is doing F[7] at t and this

[6] The reference to the 'fulfillment of the goals of the agent who is using the artifact' is by no means new in trying to understanding artifact function. Boorse's account of function (Boorse 1976), for example, takes the function of an artifact to be fixed by the 'goal-directed human activity' of what this artifact is used for: the function of fountain pens is to write because they contribute to the goal-directed human activity of writing. Note, however, that while for Boorse, the function depends on what happens in the actual world only (whether an artifact (in fact, an artifact-type) contributes, in the actual world to a goal-directed human activity), my notion of artifact function is a modal one: what matters is whether an artifact (always a token artifact) would contribute to the fulfillment of the goals of the agent who is using the artifact. It is this modal aspect of my definition that makes my account capable of handling all the three objections I raise in the next section.

[7] The phrase 'doing F' here merely means that x has some features or does something – it does not mean having the function to do so (otherwise the account would be circular. So if we find something x does that is such that possible worlds where it's doing so contributes to the fulfillment of the goals of the user are closer than possible worlds where it doesn't, we have found a function for x.

contributes to the fulfillment of the goals of the agent who is using the artifact are closer to the actual world than any of those possible worlds where x is doing F at t, but this does not contribute to the fulfillment of the goals of the agent who is using the artifact.[8]

The function of the slinky is to roll from one step to the other because some 'relatively close' possible world where it is rolling from one step to the other and this contributes to the fulfillment of the goals of the agent who is using it are closer to the actual world than any of those possible worlds where it is rolling from one step to the other, but this does not contribute to the fulfillment of the goals of the agent who is using it. What it was designed for is irrelevant.

One may be slightly suspicious of the reliance on the notion of 'the fulfillment of the goals of the agent who is using the artifact', so I need to make some explanatory remarks about this notion.

First, what if nobody is using the artifact at the moment? Would it follow that the artifact has no function? No. As we have seen, artifact function is defined by what *would* contribute to the fulfillment of the goals of the agent who is using the artifact. If nothing contributes to the fulfillment of the goals of the agent who is using the artifact in the actual world, say, because nobody is using the artifact, this does not mean that the artifact has no function. Whether it has a function depends not just on what happens in the actual world, but on what would happen if things were different. If some 'relatively close' possible world where someone is using the artifact and what it is doing contributes to the fulfillment of the goals of this agent are closer to the actual world than any of those possible worlds where someone is using the artifact and what it is doing does not contribute to the fulfillment of the goals of this agent, then it does have a function.

Second, 'the fulfillment of the goals of the agent who is using the artifact' is the equivalent of the organism's inclusive fitness in the biological case: A's trait, x, has a biological function F if x's doing F would contribute to A's inclusive fitness. Conversely, A's part, x, has an artifact function F if x's doing F would contribute to the fulfillment of the goals of the agent who is using A.

But while the inclusive fitness of an organism is a straightforward scientific concept, the notion of the goals of the agent who is using the artifact is neither straightforward, nor scientific. One may worry that if artifact function is defined in terms of what would contribute to the fulfillment of the agent who is using the artifact, then the function of an artifact depends on the intentions of the agent who is using it. This seems very different from biological function, where, as inclusive fitness is not relative to anyone's intentions, function is not either. But note that according to the standard analysis of artifact function, function also depends on intention – in this case, on the intention of the designer. If we accept the modal theory of artifact function, we get a very similar picture. Function depends on intention: not on the intention of the designer, but on that of the user of the artifact.

The conclusion is that the function of artifacts, like the function of biological traits, is fixed by modal facts. What they were designed for is not all that important. Before turning to some objections, we need to make sure that the modal theory of function can account for the difference between not having a function and malfunctioning.

[8] Again, as in the case of biological function, here x may or may not do F in the actual world, but when assessing what worlds are closer, we should ignore the actual world.

Any theory of function needs to be able to explain malfunctioning. As we have seen, an artifact malfunctions if and only if it has a function, but it fails to perform it. This is perfectly possible in my account, since even if x is not doing F in the actual world, it may still be true that if x were performing F, then this would contribute to the fulfillment of the goals of the agent who is using x (or the object x is a part of). If an artifact is not doing (or even cannot do) F in the actual world, say, because it is broken, but in a 'relatively close' possible world it is doing F and its doing F contributes to the fulfillment of the goals of the agent who is using it, then we can still attribute function F to it. This is exactly what happens if a trait is malfunctioning: if it fails to perform its function.

7 Objections

I need to address a couple of possible objections.

7.1 First Objection: Function and Use

The first, and most important, objection concerns the distinction between the function of an artifact and what this artifact could be used for. I could use my laptop as a doorstop, but does this make it the (or, at least, a) function of the laptop to serve as a doorstop? One important worry about the account of artifact function I outlined here is that it makes the notion of x's function dangerously similar to what x could be used for. And every given object could be used for thousands of things. Thus, the danger is that the account I have been proposing proliferates function: every object will end up having thousands of function.

It is important to point out that if I use an object as a doorstop in the actual world, then it does not follow under my definition that a function of this object would be to be a doorstop. Again, the function of artifacts is fixed by counterfactual facts. The function of an artifact is not whatever it does that fulfills the goals of the agent who is using it, but what it does that *would* contribute to the fulfillment of the goals of this agent.

Suppose that while moving into my new apartment, I put my laptop on the floor in order to keep a door open. Does it have the function to serve as a doorstop? Not really. It fulfills my goal: I want it to keep the door open and it does keep the door open. Now the question we should ask is whether its serving as a doorstop *would* contribute to the fulfillment of my goals with it. If things were a little different, would the laptop's serving as a doorstop contribute to the fulfillment of my goals with it? The answer depends on how different things could be.

In the scenario I described, there is little modal flexibility: we do not have to go very far from the actual world to find a world where my computer serves as a doorstop, but this does not contribute to the fulfillment of my goals with it. But I may like this arrangement and keep my computer installed against the door, in which case, we would have more modal flexibility: in nearby possibly worlds, the computer's serving as a doorstop contributes to the fulfillment of my goals with it.

Thus, if a function of an artifact x is to do F, it is not enough that I happen to use x for F-ing in the actual world (or that I could do so). If it is true that function attribution

has modal force, then it also needs to be true that if things were different, x would still be used for F-ing.

7.2 Second Objection: Function and Usefulness

A related worry is that, with or without the modal spin, my definition does not capture the notion of function, but rather the notion of usefulness. My response is to bite the bullet: function may have a lot to do with usefulness.

The main consideration against thinking about function as usefulness is that the notion of function is generally taken to be tied to the notion of design. But if the argument I presented above is correct, then this strategy of tying function to design is misguided as, at least in the biological case, what was being designed (or selected for) are different tokens of the same type as the token whose design/function we are explaining. But if function is importantly different from design, then we have no prima facie reason for rejecting the link between function and usefulness.

Further, it has been argued recently that if we conceive of function as usefulness we may avoid some undesirable consequences of conceiving of function as design (Christensen and Bickhard in preparation; Christensen et al. in preparation). And here I need to drag in the thorny issue of normativity. Function is a normative concept and so is usefulness. But it is not clear that design, if we strip the concept from its traditional close association with function, is normative (Cameron 2004). The modal theory of artifact function can account for the normativity of function with the help of the modal facts that fix this function (see Nanay 2010). But it is not clear that the normativity of function can be explained with reference to the past (see, again, Cameron 2004).

8 Conclusion: Artifact Function and Artifact Categorization

I argued that artifact function is not determined by design. Instead of exporting a standard account of function from the domain of artifact to the biological domain, we should follow a different strategy. We should import a modal account of function from the biological to the artifact cases.

Most of the philosophical literature on function has been focusing on biological function neglecting the analysis of the attribution of artifact function. Many of the most influential attempts to characterize biological function take an implicit account of artifact function-attributions for granted and use considerations about what is assumed about artifact function to arrive at a viable theory of biological function.

I followed the opposite route. I did not take anything for granted about artifact function and proceeded to use independent considerations about biological function in order to give an account of artifact function. Biological function is not determined by past selection, but by modal facts. And artifact function is also determined by modal facts. Thus, we can restore the symmetry between artifact and biological function-attributions.

Crucially, if we accept the modal theory of artifact function, we can resolve the debate about artifact categorization. The empirical findings about artifact categorization show that we individuate artifact categories depending on our practical interests. Artifact categorization is context sensitive. If we accept the modal theory of artifact

function, according to which function attribution is context sensitive and it depends on our practical interests, there is no impediment to claiming that we individuate categories by means of their function. It is just that the function by means of which we individuate artifact categories also is as dependent on our practical interests as the artifact categories themselves.

In short, we can reconcile the view that artifact categories are individuated in functional terms with the view that artifact categorization is sensitive to context and to our pragmatic interests, as long as we use the modal theory of artifact functions. As I aimed to show, there are other reasons for accepting the modal theory of artifact function – for example, to maintain the symmetry with biological functions. But if accepting this theory helps us to reconcile seemingly contradicting views about artifact categorization, this in itself may be a good reason for accepting it.

Acknowledgements This work was supported by the EU FP7 CIG grant PCIG09-GA-2011-293818 and the FWO Odysseus grant G.0020.12N.

References

Ahn, W. 1998. Why are different features central for natural kinds and artifacts? The role of causal status in determining feature centrality. *Cognition* 69: 135–178.

Ahn, W.-K., C. Kalish, S.A. Gelman, D.L. Medin, C. Luhmann, S. Atran, J.D. Coley, and P. Shafto. 2001. Why essences are essential in the psychology of concepts. *Cognition* 82: 59–69.

Allen, C., and M. Bekoff. 1995. Biological function, adaptation, and natural design. *Philosophy of Science* 62: 609–622.

Beckner, M. 1959. *The biological way of thought*. New York: Columbia University Press.

Bigelow, J., and R. Pargetter. 1987. Functions. *Journal of Philosophy* 84: 181–197.

Bloom, P. 1996. Intention, history, and artifact concepts. *Cognition* 60: 1–29.

Bloom, P. 1998. Theories of artifact categorization. *Cognition* 66: 87–93.

Bloom, P. 2007. More than words. *Cognition* 105: 649–655.

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

Braddon-Mitchel, D., and F. Jackson. 2002. A pyrrhic victory for teleonomy. *Australasian Journal of Philosophy* 80: 372–377.

Buller, D.J. 2002. Function and design revisited. In *Functions*, ed. A. Ariew, R. Cummins, and M. Perlman, 222–243. Oxford: Oxford University Press.

Burge, T. 1989. Individuation and causation in psychology. *Pacific Philosophical Quarterly* 70: 303–322.

Cameron, R. 2004. How to be a realist about *sui generis* teleology yet feel at home in the 21st century. *The Monist* 87: 72–95.

Chaigneau, S.E., L.W. Barsalou, and S.A. Sloman. 2004. Assessing avordance and intention in the HIPE theory of function. *Journal of Experimental Psychology. General* 133: 601–625.

Christensen, W. D. Bickhard, M. H. in preparation. *Function as Design vs. Function as Usefulness*.

Christensen, W.D., and M.H. Bickhard. in preparation. The process dynamics of normative function. *The Monist* 85: 3–28.

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

Cummins, R. 2002. Neo-teleology. In *Functions*, ed. A. Ariew, R. Cummins, and M. Perlman, 157–173. Oxford: Oxford University Press.

Davies, P.S. 2000. Malfunctions. *Biology and Philosophy* 15: 19–28.

Davies, P.S. 2001. *Norms of nature: Naturalism and the nature of functions*. Cambridge: The MIT Press.

Dennett, D.C. 1996. Cow-sharks, magnets and swampmen. *Mind and Language* 11: 76–77.

Dipert, R.R. 1993. *Artifacts, art works, and agency*. Philadelphia: Temple University Press.

Godfrey-Smith, Peter. 1994. A modern history theory of functions. *Nous* 28: 344–362.

Griffiths, P.E. 1993. Functional analysis and proper functions. *The British Journal for the Philosophy of Science* 44: 409–422.
Houkes, W., and P.E. Vermaas. 2004. Actions versus functions: A plea for an alternative metaphysics of artifacts. *The Monist* 87: 52–71.
Kitcher, Philip 1993 Function and Design *Midwest Studies in Philosophy 18*: 379–397. Reprinted in David L. Hull and Michael Ruse (eds.): *The Philosophy of Biology*. Oxford: Oxford University Press. pp. 258–279.
Lewens, T. 2004. *Organisms and artifacts*. Cambridge: The MIT Press.
Lewis, D. 1973. *Counterfactuals*. London: Blackwell.
Malt, B.C. 1994. Water is not H2O. *Cognitive Psychology* 27: 41–70.
Malt, B.C. 1995. Category coherence in cross-cultural perspective. *Cognitive Psychology* 29: 85–148.
Malt, B.C., and E.C. Johnson. 1992. Do artifact concepts have cores? *Journal of Memory and Language* 31: 195–217.
Malt, B.C., and S.A. Sloman. 2007. Category essence or essentially pragmatic? Creator's intention in naming and what's really what. *Cognition* 105: 615–648.
Millikan, R.G. 1984. *Language, thought and other biological categories*. Cambridge: MIT Press.
Millikan, R.G. 1996. On swampkinds. *Mind and Language* 11: 103–117.
Mills, S., and J. Beatty. 1979. The propensity interpretation of fitness. *Philosophy of Science* 46: 263–286.
Nanay, B. 2006. Symmetry between the intentionality of minds and machines? The biological plausibility of dennett's account. *Minds and Machines* 16: 57–71.
Nanay, B. 2010. A modal theory of function. *Journal of Philosophy* 107: 412–431.
Nanay, B. 2011. Function, modality and mental content. *Journal of Mind and Behavior* 32: 84–87.
Nanay, Bence 2012 Function attribution depends on the explanatory context: Reply to Neander and Rosenberg's reply to Nanay. *Journal of Philosophy* 109: 623–627.
Neander, K. 1991a. Functions as selected effects. *Philosophy of Science* 58: 168–184.
Neander, K. 1991b. The teleological notion of 'function'. *Australasian Journal of Philosophy* 69: 454–468.
Neander, K. 1996. Swampman meets swampcow. *Mind and Language* 11: 118–129.
Neander, K. 2002. Types of traits. Function, structure, and homology in the classification of traits. In *Functions*, ed. A. Ariew, R. Cummins, and M. Perlman, 402–422. Oxford: Oxford University Press.
Neander, Karen and Alex Rosenberg 2012 Solving the circularity problem for functions: a response to Nanay. *Journal of Philosophy* 109: 613–622.
Papineau, D. 1996. Doubtful intuitions. *Mind and Language* 11: 130–132.
Papineau, D. 2001. The status of teleosemantics, or how to stop worrying about swampman. *Australasian Journal of Philosophy* 79: 279–289.
Preston, B. 1998. Why is a wing like a spoon? A pluralist theory of function. *Journal of Philosophy* 95: 215–254.
Rips, L.J. 1989. Similarity, typicality, and categorization. In *Similarity and analogical reasoning*, ed. S. Vosniadou and A. Ortony, 21–59. New York: Cambridge University Press.
Sloman, S.A., and B.C. Malt. 2003. Artifacts are not ascribed essences, nor are they treated as belonging to kinds. *Language & Cognitive Processes* 18: 563–582.
Vermaas, P.E., and W. Houkes. 2003. Ascribing functions to technical artefacts: a challenge to etiological accounts of functions. *The British Journal for the Philosophy of Science* 54: 261–289.
Walsh, D.M. 1996. Fitness and function. *The British Journal for the Philosophy of Science* 47: 553–574.
Williams, G.C. 1966. *Adaptation and natural selection*. Princeton: Princeton University Press.
Wright, L. 1973. Functions. *Philosophical Review* 82: 139–168.