

# Perception, History and Benefit

(2015, *Episteme*, 13 (1), 61-76)

Mona Simion<sup>1</sup>

## 1. Introduction

On several recent views in the literature, history matters for epistemic entitlement<sup>2</sup> (e.g. Millikan (1984b), Lyons (2009), Graham (2012, 2014)). According to Peter Graham, for instance, epistemic warrant turns on the etiological function of our belief-acquisition mechanisms, which, in turn, turns on histories of positive biological benefit that explain why the relevant mechanisms exist or operate the way they do.

This paper aims at putting pressure on this claim. It is argued that biological benefit-based historical explanations of why epistemic purposes arise, and, in turn, of how the norms that are associated with them come into being, are irrelevant to the content of the relevant epistemic norms. Inasmuch as we take epistemic entitlement to reduce to propriety by the epistemic norm,<sup>3</sup> they will not affect the epistemic status of our beliefs either.

For the sake of brevity, I will restrict the discussion to Graham's proper functionalist view; however, the results of this paper generalize to all attempts to explain epistemic normativity by employing benefit-based etiological accounts of function acquisition. I will proceed as follows: in §2 I briefly outline Graham's account, together with some recent refinements. In §3 I put forth worries with regard to both the central necessity claim and the sufficiency claim involved in his proper functionalism; I argue that function acquisition via positive biological feedback is irrelevant for epistemic entitlement. In the last section (§4) I go on a rescue mission on behalf of Graham and show how his account can be modified in order to accommodate the results of this paper. In §5 I conclude.

---

<sup>1</sup> University of St Andrews and KU Leuven Epistemology Group, Andreas Vesaliusstraat 2, Leuven 3000, Belgium; [mis2@st-andrews.ac.uk](mailto:mis2@st-andrews.ac.uk); [mona.simion@hiw.kuleuven.be](mailto:mona.simion@hiw.kuleuven.be); [www.mona-simion.com](http://www.mona-simion.com)

<sup>2</sup> For the purposes of this paper, I will use entitlement and justification interchangeably. Also, here and below, what is discussed is *prima facie, pro tanto* justification.

<sup>3</sup> Which Graham does and this paper accepts.

## 2. Graham on Entitlement

Epistemic entitlement, according to Graham, consists in the proper performance of belief-forming processes that have reliably forming true beliefs as an etiological function (henceforth, e-function) (Graham, 2012, 463). On the etiological theory of functions,<sup>4</sup> functions turn on histories that explain why the item exists or operates the way it does.

### 2.1 *Proper Function*

Roughly, the thought goes as follows: our belief-acquisition processes have been selected for their reliability in producing true beliefs. Because they reliably produced true beliefs in our ancestors, they enabled them to survive – spot dangerous predators, find food and so on – and reproduce, thereby replicating the same cognitive systems with the same function in their descendants. As such, our belief-acquisition processes acquired the e-function of reliably producing true beliefs. The etiological account is an account of functions as purposes; by being selected for it, our belief-acquisition processes have acquired the purpose of reliably producing true beliefs.

But purposes also come with associated norms, which regulate the correct procedure to follow in order to reach them. Epistemic entitlement consists in meeting epistemic norms. Graham has a naturalistic story to tell about this too. Given its e-function, the process will form true beliefs reliably in normal conditions, i.e., conditions similar to those in which it was selected, when functioning normally, that is, when functioning in the way in which it was functioning when it was selected for its beneficial effects. The standards for normal functioning are thus both natural and epistemic norms; they are constitutively associated with both natural selection and promoting true beliefs and avoiding error. When functioning normally – whether in normal conditions or not – the belief-forming process will fulfil the procedural norm constitutively associated with forming true beliefs reliably; it will work the way it is supposed to work, where the right way of working is partly constituted by reliably getting things right in normal conditions. As such, according to Graham, *de facto* reliability does not matter for epistemic entitlement; reliability in certain circumstances – normal circumstances – matters (2012, 466).

---

<sup>4</sup> Defended by people like Larry Wright (1976), Ruth Millikan (1984a, 1984b), David J. Buller (1998).

Now, Graham's account has quite a lot going for it. First and foremost, it offers a very plausible naturalistic story explaining the existence of such oddities like goals and norms; also, with regard to epistemic goals in particular, it accounts for why nature should care about things like truth and knowledge in the first place.

Furthermore, the account is meant to solve two of the most notable challenges associated with classical process reliabilism; namely The New Evil Demon problem<sup>5</sup> and the case of Norman, the clairvoyant.

This paper will focus on Graham's proper functionalist solution to the latter. Here is the famous case by Laurence Bonjour:

NORMAN: Norman, under certain conditions which usually obtain, is a completely reliable clairvoyant with respect to certain kinds of subject matter. He possesses no evidence or reasons of any kind for or against the general possibility of such a cognitive power or for or against the thesis that he possesses it. One day Norman comes to believe that the President is in New York City, though he has no evidence either for or against this belief. In fact the belief is true and results from his clairvoyant power under circumstances in which it is completely reliable (Bonjour 1985, 41).

Classical process reliabilism,<sup>6</sup> roughly put, claims that one is justified in holding a belief if and only if one has acquired it via a reliable process – that is, a process that delivers beliefs with a favourable truth to falsity ratio. The case of Norman sheds doubt on the sufficiency claim. Norman's clairvoyance is, by stipulation, a very reliable process; however, we are not inclined to think that Norman is justified in his belief that the President is in New York.

At first glance, Graham's account does not seem to share this weakness with classical process reliabilism. Recall that, in Graham's view, not just any reliable process will do for entitlement, it also needs to be the case that the process used comes with the right type of history; this is where the etiological constraint steps in. Norman's clairvoyance has not been selected for the function of reliably producing

---

<sup>5</sup> This case is an attack on the reliabilist necessity claim and has been put forth in Lehrer and Cohen (1983). Suppose that an evil scientist kidnaps Philip, a normal adult, and hooks his brain up to a super computer that is programmed to induce the kind of sensory experiences he would have had if he had not been kidnapped. The thought is that, intuitively, Philip is as justified after envattment as he was beforehand, although, in fact, his cognitive processes hardly ever deliver truths. If that is the case, de facto reliability of the processes used seems to be too strong a condition on justification.

Graham's account has no problem accommodating the intuition of justification here; after all, envatted Philip does use cognitive processes that 1) have the function of reliably delivering true beliefs – acquired before Philip was kidnapped –, and 2) are functioning properly, i.e. in the way in which they did back in the day when they acquired their function. Similarly to Philip's cognitive system, a heart placed in a container with nutrients is still a properly functioning heart, which only fails to fulfil its function due to the change in environment.

<sup>6</sup> Most notably associated with Alvin Goldman (1979).

true beliefs; as such, it has no function, so nothing counts as functioning normally – it is not subject to any epistemic norms regulating its proper function – and, consequently, cannot confer entitlement on Norman’s beliefs.

## 2.2 *How Much History?*

In spite of the great intuitive appeal of Graham’s naturalistic picture, several authors<sup>7</sup> have expressed worries with regard to the historical constraint on epistemic entitlement being unmotivated.

Recall that if epistemic entitlement turns on functions, and functions turn on history, then entitlement requires history. However, take, for instance, Davidson’s Swampman (1987, 443). Swampman comes into existence as a result of a lightning hitting a swamp and strangely rearranging gas molecules into the exact duplicate of some ordinary person. By being created as a result of a completely random process, Swampman’s cognitive processes lack a function, thus cannot be functioning properly or improperly. As a result, on Graham’s account, he cannot have justified beliefs. However, intuitively, at least after acquiring the necessary concepts, Swampman’s belief that he is sitting in a swamp seems perfectly justified (Sosa 1993).

Keep in mind that the case of Swampman is actually not as odd as it might sound. Graham himself notes that “Swampman is just a full-body example of what biologists call hopeful monsters. [...] A hopeful monster is a beneficial macro-mutation” (2014, 30). Given that we are interested in Swampman’s epistemic status, Sosa’s argument does not rest on the metaphysical possibility of Davidson’s full-body example. All Sosa needs is a perfectly normal person with an epistemically beneficial macro-mutation, in the shape of a new belief-acquisition system.<sup>8</sup>

In response to this, Graham (2014) goes more into detail about what the etiological constraint consists in. According to him, while etiology does require some history, it does not require an awful lot of it; there are several ways to cash out the etiological requirement which do not presuppose directional selection, i.e. selection over generations. A trait can also acquire a particular function by on-going, maintenance selection, or through a learning process, or even by the metabolic activity of the organism itself. We will look more closely to what all these alternative ways to acquire e-functions consist in later in the paper. For now, the rough picture will suffice: what it all amounts to, eventually, is explaining the existence/continuous

---

<sup>7</sup> E.g. Sosa (1993), Evans and Smith (2012), Ghijzen (forthcoming).

<sup>8</sup> As such, here and below, sceptics about the metaphysical possibility of a full-body Swampman coming into existence should read all talk of Swampman as involving a normal individual endowed with a more limited mutation, restricted to her cognitive mechanisms.

existence of a trait through a through a longer or shorter history of positive biological feedback which may or may not be selective:

Functions arise from consequence etiologies, etiologies that explain why something exists or continues to exist in terms of its consequences, because of a feedback mechanism that takes consequences as input and causes or sustains the item as output (Graham 2014, 35).

Of course, Swampman's perceptual system will not have a function the first time it gets employed. Nor will his heart, for that matter. Only one cycle is not enough for function acquisition, and thus there is no such thing as proper functioning either. However, as soon as the positive feedback mechanism is in play, Swampman is going to enjoy as much justification as the rest of us. The same, of course, will go for Norman: as soon as his clairvoyant beliefs benefit his survival, thereby sustaining Norman's clairvoyant ability, Norman will come to enjoy justification.

Let us take stock: we have seen that, according to Graham, the following two conditions are individually necessary and jointly sufficient for epistemic entitlement:

(i) *The Function Acquisition Requirement (FAR)*: The cognitive mechanism enjoys a (longer or shorter) history of (selection and) positive biological feedback.

(ii) *The Normal Functioning Requirement (NFR)*: The cognitive mechanism functions in the way it did back when it acquired its function.

Now, what I am going to do in the following section is present worries with respect to both the FAR necessity claim and the FAR + NFR sufficiency claim involved in Graham's proper functionalism. As such, I will first look at the role the e-functional explanation of purpose acquisition is supposed to play in fixing the content of the relevant epistemic norm. I will argue that FAR, even when spelled out in terms of the weakest etiological explanation of the market, is not necessary for epistemic entitlement. Further on, I will target Graham's sufficiency claim. I conclude that a history of positive biological feedback, no matter how short, is completely irrelevant to the content of the epistemic norm.

### **3. The Irrelevance of Biological Benefit**

Here is the first worry that motivates this paper. Norms are associated with reaching goals. If I want to take the shortest route from point A to point B, there is going to be (at least) a norm stipulating how I should go about it, when to turn right, when to

turn left, etc. This claim need not imply a commitment to a consequentialist value-theoretic framework. All I need for the following argument to go through is that norms are, somehow, *associated* with the relevant goals, no matter which explains the other. The consequentialist will explain the ‘ought’ in terms of the ‘good’; he will say that the norm is there to guide us in reaching the goal. As opposed to this, the deontologist reverses the order of explanation; he would have it that the goal is only valuable in virtue of the fact that the norm gives us reasons to favour it. Anyhow, the mere association claim holds.<sup>9</sup>

Now, importantly, notice that the content of the norm will be completely independent of whether reaching the goal benefits me or not. I might want to reach point B because that would earn me a million dollars, for no good reason whatsoever, or because I just want to torture myself by doing something I really hate. However, no matter the explanation of how the goal arose, the norm associated with my goal is going to have the same content: in order to reach point B via the shortest route I have to, say, take ten steps forward and ten steps to the right.

Consequently, my actions towards that goal are going to be proper or improper actions independently of why I decided to reach point B in the first place. Proper or improper with respect to reaching that particular goal, that is. Proper actions, of course, are going to obey the norm(s) associated with my goal, and thus promote my reaching it. As such, in view of my goal, and in accordance with the norm observing it, I am going to be entitled to some actions, but not to others. In view of reaching point B via the shortest route, I am entitled to take ten steps forward and ten steps to the right, but not to take ten steps back or three steps to the left.<sup>10</sup> All this is independent of whether reaching point B benefits me or not.

In the light of this, this section is going to argue that FAR is false: positive biological feedback is irrelevant to epistemic justification. To this aim, what I will do first is go through all the ways in which Graham (2014) spells out the etiological constraint, from the strongest to the weakest. I will argue that each of them imposes too strong a requirement on epistemic entitlement, by bringing cases of individuals enjoying intuitively justified beliefs, although they fail to comply with the requirement in question. As such, I will show that even the weakest etiological incarnation of FAR put forth by Graham fails to be necessary for epistemic entitlement. At the end of this section, I will also look at the sufficiency claim and shed significant doubt upon it too.

Let us go back to the several ways of e-function acquisition put forth in Graham (2014). What Graham gives us is a set, W, of ways in which an e-function may arise;

---

<sup>9</sup> Thus, the argument made by this paper can be constructed in both consequentialist and deontological terms; nothing here hinges on this.

<sup>10</sup> Norms generated by a particular goal can, of course, be overridden, by other norms, pertaining to a different goal. Again, we are discussing *prima facie* propriety and entitlement.

that is, different ways to spell out the history of positive biological feedback requirement (FAR). What I will do next is look at them in turn and argue that, for each member,  $w$ , of  $W$ , an agent's belief-forming processes may produce justified beliefs in absence of a history of selection and/or positive biological feedback. Since justified beliefs satisfy the epistemic norm, for all  $w$  in  $W$ , satisfying the epistemic norm does not require a history of selection or positive biological feedback. Furthermore, I will show that an agent's belief-forming processes may even produce justified beliefs in spite of a history of *negative* biological feedback triggered by reliably producing true beliefs. As such, the point generalizes beyond  $W$ : a history of selection and/or positive biological feedback is not required for epistemic justification, and, therefore, the FAR necessity claim is false.

Let us put some order into our set first. Buller (1998) distinguishes between a strong etiological theory of functions (henceforth, ST) and a weak incarnation thereof (WT). They both require some history, and a positive feedback mechanism, as they both aim to explain why a particular trait exists or continues to exist in a system. Importantly, what sets them apart is the variation requirement. On the strong theory, for a trait to acquire a function, it must have been selected from among weaker traits for that particular function. On WT, that need not be the case. As such, what we have is the following common requirement for both ST and WT:

**(R1)** E-functions require a history of positive feedback,

while ST is also going to ask for selection:

**(R2)** There must have been variation with respect to trait  $T$  at some point in history, and  $T$ 's bearers must have had greater fitness than bearers of its variants because of  $T$ 's producing effect  $E$ .<sup>11</sup>

Now, the way in which both the selection requirement and the positive biological feedback requirement can be spelled out vary, as evidenced by Graham's own proposed ways of function acquisition. If that is the case, let us look at both of these incarnations of the etiological theories in turn, and, in the process, cover all of Graham's proposals for spelling out of the etiological constraint, in order to check the necessity claims involved here. I will, as dialectically appropriate, start with the proposed versions of ST, i.e. ways of function acquisition that require both R1 and R2, and argue that R2 is too strong: selection is not necessary for epistemic justification. Further on, I will move to discussing incarnations of WT and shed doubts on the relevance of R1, i.e. positive biological feedback, for entitlement.

---

<sup>11</sup>Adapted from Buller (1998, 512).

### *3.1 The Variation Requirement (R2)*

#### (i) Directional Natural Selection

Take, first, natural selection. One way to understand it is directionally, that is, as selecting beneficial traits over generations. After all, plausibly, our cognitive system was thus selected for the benefits delivered by its reliably producing true beliefs.

This, of course, will require quite some amount of history before a trait acquires a function. Is it necessary, however, that one's beliefs are formed via a mechanism enjoying such extended history credentials? Well, the answer seems to be negative; we have already seen that cases like Swampman show that a long history of selection that spans generations is not necessary for our belief-acquisition mechanisms to deliver epistemic justification.

#### (ii) Maintenance Natural Selection

What about short cycles of selection? One way to cash this out is in terms of maintenance selection. This latter type is recent, on-going selection:

Most mutations, for example, are harmful. [...] Because harmful they are selected out, in favour of the normal variant of the trait. The non-malformed trait then continues to exist and predominate in the population because of its relative superiority (Graham 2014, 27).

On this account, as opposed to harmful traits, your reliable belief-acquisition mechanisms benefit you, thereby explain your continuous existence, and thereby explain their continuous existence. Like in the case on beneficial traits being maintained versus mutations being selected out, Swampman's reliable belief acquisition system will continue to exist due to its being more beneficial for Swampman's survival than systems which do not perform as well to their holders' benefit.

However, this way of spelling out the variation requirement looks absurd; it hardly seems to be the case that Swampman cannot ever acquire justified beliefs unless there are other creatures around, endowed with less beneficial belief-acquisition mechanisms. Nor does it look plausible that epistemic entitlement requires that Swampman must have been initially endowed with weaker traits too, so that the one he is using now got selected for being the most beneficial. If there are no weaker systems around, however, there is no maintenance selection.



### (iii) Trial and Error Learning

Another, more plausible way to cash out the selection requirement would be in terms of trial and error learning; in a sense, the latter can be seen as a 'short history' version of natural selection (Graham 2014, 33). They both require variation, and they both select in the beneficial traits/behaviours and out the ones that deliver negative feedback. Thus, instead of nature selecting in successful traits and selecting out less successful ones, maybe what is needed for function acquisition and, thereby, entitlement, is that Swampman learns how to use his perception ability through trying several variations in behaviour and selecting in the successful ones over the ones that deliver negative feedback. After all, that is how, for instance, my fingers acquired, among other functions, the function to tie my shoelaces: I've tried a number of times, and selected in the best ways to go about it.

Notice, though, that selection looks like too strong a requirement again. After all, the case stipulates that Swampman's perceptual faculties are going to reliably deliver true beliefs. As such, Swampman's learning is not going to involve any variation: Swampman will not be in a position to choose the more reliable ways to go about belief acquisition over the less reliable ones. His new belief acquisition mechanism is just going to work perfectly fine in reliably delivering true beliefs. Still, it looks as if after a few successful occurrences, the mechanism is going to provide him with justification.

## *3.2 The Positive Biological Feedback Requirement (R2)*

So maybe the problem is that R1 (the variation requirement) is too strong. Let us then turn towards the weak incarnation of the etiological theory and see how it deals with the epistemic cases. Recall that according to WT, e-functions only require a (longer or shorter) history of positive feedback.

Again, the 'long history' incarnation is not going to work better than its stronger counterpart when it comes to recently acquired mechanisms like Swampman's. Let us then look at the available short-history ways of spelling out this requirement:

### (i) Metabolism

For the short history version, one promising candidate for function acquisition is the normal operation of the metabolism<sup>12</sup> of an organism, which only requires that the existence or continued existence of a trait is explained in terms of its consequences, because of a feedback mechanism that takes the beneficial consequences as input and causes or sustains the item that produced them as output. Your heart, for instance, by pumping blood, keeps you alive and contributes to its own continued existence:

Where natural selection takes earlier generations [...] as input, the normal operation of the metabolism of an organism takes earlier cycles as input [...]; it explains the continued existence of traits within individuals over *cycles* in virtue of their beneficial effects [...]. An earlier *cycle* of your heart, by pumping blood, contributes to the existence of a later *cycle*, partly in virtue of its beneficial effects to you (Graham 2014, 29).

As such, just by keeping itself alive through this positive feedback mechanism, the heart acquires the function of pumping blood. This already looks much more plausible than what we have seen so far. Think of Swampman: plausibly, with every new true belief, he gets more information about his environment, and thereby manages to do a better job at adapting to it and staying alive. Consider, also, Norman: if every time he forms the clairvoyant belief that there is free food offered at the market, he goes there and gets free food, this will benefit his survival, which, in turn, will (at least partly) explain<sup>12</sup> why the clairvoyance ability is still there. Also, after a short while, it surely looks like he is justified in his clairvoyant beliefs.

Is it the case that positive biological feedback is necessary for justification, however? Consider the following case:

SICK SWAMPMAN: Swampman suffers from a terminal disease due to his weird way of having come into being. The radiations that helped with his birth also resulted in his having a large malignant lung tumour that is likely to kill him in three months. As with many cancer patients, Swampman is better off not believing that he is going to die soon; this would just destroy his morale and quicken his death. As opposed to this, blissful ignorance might prolong his life substantially. Now, Swampman hears his doctors saying that he is going to die in three months' time. Furthermore, say that he ends up in the hospital pretty soon after having come into existence, such that most of his true beliefs will be about his unfortunate predicament, and will, therefore, worsen his situation.

---

<sup>12</sup> In a recent talk (xxxxx, 2014), Graham himself restricts his etiological account to metabolism as a source of function acquisition.

Swampman's belief-acquisition systems are not benefiting him in any way; again, though, it looks as if, as soon as he acquires the relevant concepts, Swampman is perfectly justified in believing he suffers from a terrible disease.<sup>13</sup> Thus, again, this way of spelling out the etiological constraint for function acquisition seems to fail to be necessary for epistemic justification. Furthermore, it turns out that most of the true beliefs produced by one's cognitive system can be harmful, while one is still perfectly justified in holding them.

## (ii) One-Off Learning

In the light of all this, all Graham is left with for spelling out the etiological constraint on function acquisition is one-off learning. The latter is a limit case of trial and error learning, where I learn how to do something without selecting out any errors. Although no selection is going on in this case, according to Graham, one-off learning is a source of indirect function acquisition. Here is how: the innate capacity to learn has the function of helping the creature adapt to its environment, which, in turn, generates positive feedback by benefitting the creature's survival. As such, one-off learning to X inherits this function indirectly (2014, 34).

Presumably, however, in order for this to be the case, the capacity to learn must itself have acquired the function of facilitating adaptation in some way to begin with. But, in the case of Swampman, it can't be natural selection or maintenance selection, for reasons already mentioned. The short history proposals also don't seem to work. After all, we may assume that Swampman never had occasion to exercise this capacity before, and he still seems to be able to enjoy justified beliefs.

Notice also that we can simply stipulate that Swampman's learning abilities are generally hurting him, due to some evil genius who makes sure that, every time Swampman learns something new, he gets a terrible disease. If learning is not biologically good for Swampman, though, it has no e-function; therefore, there is no e-function to be inherited by his belief-acquisition processes either. However, he still seems perfectly justified in believing that he is sitting in a swamp.

We have thus gone through all the available incarnations of both the strong and the weak etiological theory of functions, and there is no good news for FAR: it becomes clear that neither selection nor a history of positive biological feedback are necessary for epistemic justification. Furthermore, one's cognitive mechanisms' reliably producing true beliefs might even deliver overall negative biological

---

<sup>13</sup> Notice that switching from benefit for the individual to benefit for the species would not help save the account either. To see this, recall that the argument made by this paper can be reconstructed by replacing Swampman with a normal individual endowed with a more limited mutation, restricted to her cognitive mechanisms. In this case, this particular cognitive mechanism has never existed in any members of the species before.

feedback, without the epistemic status of the relevant beliefs being affected in any way.

### *3.3 The Sufficiency Claim*

So far, we have seen that, no matter how weak we want our etiological account to be, the necessity claim involved in FAR is still going to be too strong a requirement for epistemic entitlement. That is, we have seen that neither a history of selection, nor a history of positive biological feedback is necessary for one's cognitive mechanisms to deliver epistemically justified beliefs.

Now, let us also have a look at the sufficiency claim involved in Graham's account. Recall that, according to Graham, all that is needed for epistemic entitlement is that one's cognitive mechanisms (1) have the e-function of reliably producing true beliefs (FAR) and (2) function normally, i.e. as they did at the time of function acquisition (NFR).

As dialectically proper, let us start by first looking at the weak incarnation of the e-functions account; that is, let us see if WT, when plugged into Graham's view, is enough for epistemic entitlement. Consider the following case:

VORMAN: Vorman is a mentally challenged person living in a health institution; among other manifestations of his mental disease, Vorman also hears voices telling him stories about the whereabouts of the President all the time. Every time Vorman\* hears voices and thereby comes to form a belief about the whereabouts of the President, nurse Olga, who finds this very amusing, gives him a cookie.

Now say that, unbeknownst to both Vorman and Olga, luckily enough, Vorman's beliefs are true: the voices are part of a highly reliable clairvoyance process and that has only ever produced true beliefs. Nevertheless, it seems pretty clear that, no matter how much Olga's cookies are going to benefit Vorman biologically, he is not going to be epistemically justified in his beliefs sourced in hearing voices.

But maybe it is WT that is problematic here; that is, maybe a short history of positive biological feedback is not enough for providing the mechanism with a function. Let us look at ST then; here the claim amounts to the following: if one's cognitive system was selected for forming true beliefs due to the fact that this benefitted one's ancestors, one's beliefs are going to be justified when formed via the respective cognitive system, provided that it functions properly, i.e. the way it functioned back in the day when it acquired its function.

Here is, however, a variation on a case by Ghijsen (forthcoming) that sheds significant doubt on this way of spelling out the sufficiency claim also:

NORBERT: Norbert is the son of a mother and father who both have reliable clairvoyant abilities. As opposed to their less endowed neighbours, all Norbert's ancestors have been able to reproduce because of the benefit these clairvoyant abilities have provided for them. The abilities usually become active quite suddenly at some point after puberty. Norbert has no evidence for or against their existence in general or his own having them. At some point after puberty, Norman suddenly experiences a brief visual image of the President being in New York City and on that basis comes to believe that the president is in New York City.

Norbert seems to meet all the conditions imposed even by the ST reading of Graham's sufficiency claim – i.e., selection, positive biological feedback and normal functioning – but his beliefs still appear to lack justification. Norbert's clairvoyant faculty has the etiological function of producing true beliefs reliably, because it has the right type of history: it has been selected because it has provided a certain benefit to his ancestors. It is also perfectly reliable and functioning like it did back when it acquired its function. Still, Norbert seems to lack any justification whatsoever to believe that the President is in New York.

If that is the case, however, Graham's proper functionalist account seems to be in serious trouble: on no available reading of the e-functionalist FAR constraint, does it turn out to be either necessary or sufficient – in conjunction with NFR – for epistemic justification.

#### **4. Diagnosis**

Let us take stock and try to figure out what went wrong with Graham's account. Recall the parallel with the function of the heart: roughly, on the most modest incarnation of the e-functions account, the heart pumps blood, thereby keeping the host organism alive, and thereby continuing to exist itself. Now, we have seen that, while this explanation looks quite neat when it comes to explaining biological purposefulness and normativity, the positive biological feedback loop got us in trouble with epistemic normativity. It looks as if a cognitive system's history of biological benefit is irrelevant to epistemic entitlement. If that is the case, we are left with two options: either we (i) conclude that e-function acquisition is irrelevant to epistemic entitlement in particular, and therefore abandon the application of an e-

functionalist framework to epistemology altogether, or (ii) conclude that it is positive *biological* feedback that is not necessary for *epistemic* e-function acquisition, and attempt a rescue mission on behalf of the functionalist by offering a new way to apply the etiological framework to epistemic normativity.

What I will try to do next is look into the prospects of building a case for (ii). That is, I am going to test the following hypothesis: what we are faced with is but an application defect; biological benefit explains biological e-function acquisition, but is irrelevant to e-function acquisition outside biology. What does make a difference is domain-specific benefit.

#### *4.2 The Result Generalizes*

In the light of the above negative results, one question comes to mind: does the biological benefit constraint attached to etiological accounts of functions score equally bad when it comes to explaining normativity in other domains outside biology? That is to say, does biological benefit, although essential for biological e-function acquisition and normativity within biology, turn out to be irrelevant to other types of normativity and e-function acquisition? After all, if that is the case, maybe what we have been facing is but a symptom of a generalized application defect of the etiological theory of functions to normativity outside biology; as such, instead of abandoning the project of offering a general e-functional explanation of normativity, maybe all that is needed is a bit of work on the terms of the application procedure.

Without claiming to even come close to settling the issue, in what follows I will attempt to briefly gesture in the direction of an affirmative answer to this question. Consider, for instance, biological benefit-based teleological theories of mental content, à la Ruth Millikan (1984a) and David Papineau (1993). According to these accounts, what a representation represents depends on the functions of the system that produces the representation. The brain is a biological organ selected for information processing – due to the fact that it did that in our ancestors and they benefitted from this, which explains the continuous existence of brains – and this makes information processing one of the biological functions of the brain. As such, teleological theorists claim to account for the normative nature of mental representation: when in normal conditions – conditions of function acquisition – accurate representation stands for normal functioning, while misrepresentation is a case of malfunctioning. Here is, however, a case by Pietroski (1992) that can be read to raise similar worries as the ones raised by this paper:

Suppose a system has an innate disposition, in virtue of which it reliably produces s-tokens when conditions C, and s'-tokens when conditions C' obtain; but suppose there is no history of selective advantage in having the capacity to distinguish C from C'. Millikan's theory cannot explain any semantic difference between s-tokens and s'-tokens. But humans surely have discriminatory capacities that have never been selectively advantageous, in terms of either biological evolution or reinforced learning (1992, 276-77).

To see this, consider a perfect pitch. Say that Jerry's capacity to discriminate a well-tuned A from a slightly flat one has no history of positive biological feedback. In this case, Pietroski argues, Millikan's theory will not be able to accommodate intentional explanations like: 'Jerry winced because he noticed that the orchestra was flat'; in the absence of a history of positive biological feedback, Millikan's theory sees no difference between a well-tuned A and a flat one. (1992: 277).

One response available to the defender of teleological accounts of content would go as follows: our discriminatory capacities have been selected for having biologically benefitted our ancestors. Therefore, just like instances of one-off learning inherit the indirect function of helping us adapt to our environment from our general capacity to learn, discriminating between C and C' acquires an indirect function from the function of our discriminatory capacities in general.

Notice, however, that this response would not work for similar reasons as in the parallel epistemic case. Consider:

**TORTURED SWAMPMAN:** Swampman is captured by a sadist who will torture him every time he learns something about the world and kill him if he doesn't misrepresent the world at least once per day.<sup>14</sup>

When Swampman has survived for sufficiently many days, due to the biological positive feedback mechanism in play, his information-processing traits will have misrepresenting as an e-function; therefore, normal functioning will be misrepresenting, and cases of accurate representation will be cases of malfunctioning. Still, independently of this, it looks as if Swampman is doing something right when he represents his swamp as a swamp. Similarly, even a creature for which, by stipulation, discriminatory capacities play no role in survival, can do well in discriminating between a well-tuned A and a slightly flat one. Biological benefit is just plainly irrelevant when it comes to explaining proper representation.

---

<sup>14</sup> Thanks to XXX for this and for very fruitful discussions on this subject matter.

Of course, this is a very rough sketch of an argument for the generalization hypothesis and affords serious further research efforts; again, though, all this section attempted to do was to establish the plausibility of the generalization claim. If I am right, the problems discussed by this paper are only symptoms of a more fundamental application defect; as such, the positive biological feedback requirement for e-function acquisition will turn out to be irrelevant for all normativity outside biology. I will propose next that what will make a difference will be domain-specific positive feedback.

#### *4.2 Epistemic Positive Feedback*

First, notice that if we still want a functional account of normativity, dropping the positive feedback requirement characteristic of etiological accounts of functions does not seem to work either. Here is why: say Norman goes on forming beliefs via clairvoyance about things that are far removed from his everyday life, like the whereabouts of the President, for years; however, he never becomes aware in any way of the fact that his clairvoyant beliefs are actually true, nor does he ever use them in (practical or theoretical) reasoning. They remain completely isolated. It looks as if, no matter how long the history of his clairvoyance reliably delivering true beliefs, Norman will still not be justified in holding them. To see this more clearly, recall the more intuitive, empirically-friendly case of Vorman, who hears voices telling him stories about the whereabouts of the President all the time. Again, say that luckily and unbeknownst to him his beliefs are true. Surely, even if Vorman continues hearing voices and forming the relevant beliefs for decades, this does not help their epistemic status. Thus, at least *prima facie*, it looks like some sort of positive feedback is still needed for epistemic justification.

If that is the case, one way to rescue the e-functional explanation that readily suggests itself is by modifying its mode of application outside biology, i.e. by re-describing the positive feedback at stake by going domain-specific. After all, if biological e-function acquisition requires positive biological feedback, plausibly, epistemic e-function acquisition will require positive epistemic feedback. To see how this works, contrast the following two situations:

(i) Every time Vorman hears voices and thereby comes to form a belief about the whereabouts of the President, nurse Olga, who finds this very amusing, gives him a cookie.



(ii) Every time Vorman\* hears voices and thereby comes to form a belief that there are free cigarettes offered at the market, he goes there and gets free cigarettes (and smokes them).

Again, it seems pretty clear that, no matter how much Olga's cookies are going to benefit Vorman biologically, he is not going to be epistemically justified in his beliefs sourced in hearing voices. In contrast, at least after getting it right a number of times, it looks like Vorman\* is justified in believing there are free cigarettes offered at the market, in spite of the fact that this harms him biologically.

Thus, again, positive biological feedback does not seem to matter for epistemic function acquisition. What does matter, though, is domain-specific – in this case, epistemic – positive feedback: after getting it right on several occasions, Vorman\*'s hearing voices acquires the function of producing true beliefs.

Now, here are two main possible worries regarding this proposal: first, if positive epistemic feedback is required for function acquisition, how is this still an externalistically friendly account?<sup>15</sup> After all, one of the main motivations behind externalism regarded the epistemic status of unsophisticated cognizers. Animals and children seem perfectly justified in holding their perceptual beliefs, even though their abilities to reflect on the reliability of the cognitive processes they use in forming beliefs are not likely to be very well developed.

Here is how this is not a problem for the present account: notice that the positive epistemic feedback needed for function acquisition need not be cashed out in terms of fully conscious justified belief about the reliability of the process at stake. Sure, most likely, in adult cognizers, this is the shape it is going to take most of the time. However, take for instance, Vorman\* again; sure, he only manages to gain justified beliefs after confirmation of the reliability of his belief forming process. But need he actually have fully blown conscious beliefs about this reliability? This hardly seems to be the case. A minimal level of implicit awareness that the process works seems to be enough.<sup>16</sup>

---

<sup>15</sup> Thanks to an anonymous referee for pressing me on this issue.

<sup>16</sup> A more precise cashing out of what this minimal level of awareness amounts to, while definitely needed, falls outside of both the scope and the space limitations of the present paper; as such, it will have to make the subject of further work. One more note on the matter though: when it comes to less sophisticated cognizers, one can be even less demanding; the thought goes as follows: animals and children are less well developed than we are, and so are their belief-acquisition mechanisms and, therefore, their second-order capacities. As such, it makes sense that the level of awareness needed for the relevant epistemic feedback need not be equivalent to the one required in adults. Dogs, for instance, when finding food in the exact spot when their newly acquired clairvoyance capacities told them they will find food, will not need to actually reflect on the felicity of the event and the newly discovered reliability of their clairvoyance, in order to enjoy justification in the future. However, neither will they know there is food in that spot until the credentials of the clairvoyance process go through minimal verification, that is, until they go and find food in the relevant spot at least once. As

Furthermore, going back to the worry regarding the general fit with externalist theories, notice that this account only asks for one form or another of agent-based positive epistemic feedback 1) when it comes to newly acquired cognitive abilities – after all, perception and the like will still be nicely grounded in their initial function-acquisition moment in our ancestors, and 2) only at the time of function acquisition. As such, the account’s requirements are, in fact, pretty minimal. Furthermore, it is fair to say that the added benefits justify the costs: first, the present account gets rid of the positive biological feedback requirement, thus allowing agents who do not benefit biologically from their cognitive skills to also enjoy justification. Secondly, this account scores better than other externalist alternatives in that it 1) offers a solution to clairvoyance cases that covers all newly acquired cognitive abilities and 2) preserves the nice solution to the New Evil Daemon Problem afforded by Graham’s old account.<sup>17</sup>

A different worry one might bring is that going domain-specific for positive feedback seems, at first glance, to rob e-functional accounts of their explanation of the continued existence of a trait. After all, if forming true beliefs or accurate representations need not benefit us biologically in order for the relevant traits to have the relevant function, nothing seems to be left to explain the continued existence of the trait in question.

Notice, however, that this is not such a major loss as it might seem; after all, etiological theories of functions are intended to be real-nature theories (Millikan 1984a). And, in fact, plausibly enough, true beliefs are, most likely, beneficial for our survival. So the positive biological feedback that is, in fact, at play, does explain the continued existence of the relevant traits, even though it does not affect the content of the relevant epistemic norms in any way. All it explains is why the goals and the norms are there; because their being in place benefits us. But notice that this is all we wanted to begin with: a naturalistic explanation of the existence of such oddities as norms and purposes. As such, even though biological benefit does not affect the content of the relevant norms, it does nicely explain how and why the purposes and the norms associated with them arose in the first place.

## 5. Conclusion

---

such, not only is the present account compatible with unsophisticated cognizers enjoying justification, but it also nicely explains cases put forth by several philosophers to show that animals and children seem to gain knowledge more easily than adult cognizers: children, but not adults, seem able to acquire testimonial knowledge in a ‘room full of liars’ situation (Goldberg (2008)), while animals appear to come to enjoy justified beliefs more easily in clairvoyance like cases (Ghijsen (2014)).

<sup>17</sup> In fact, arguably, the present account improves over Graham’s old one in this respect also; after all, it does not seem to be necessary for Philip to have benefitted biologically from his true beliefs before being kidnapped for him to enjoy epistemic entitlement either in or outside the vat.

I have argued, against Graham, that biological benefit-based historical explanations of how or why epistemic goals and norms came about have no impact on the content of the relevant norms. Epistemic entitlement does not turn on a history of positive biological feedback. Rather, I have suggested, epistemic function acquisition rests on positive epistemic feedback. I have proposed that, just as biological function acquisition – and therefore biological normativity - requires positive biological feedback, epistemic function acquisition and epistemic normativity require domain-specific, that is, positive epistemic feedback. Also, I have gestured in the direction of this result generalizing for other types of normativity too. As such, if this is right, Graham's account turns out to have suffered from a defect of application of the etiological theory of biological functions to normativity outside biology.

## References

- Buller, D. J. (1998). Etiological Theories of Function: A Geographical Survey. *Biology and Philosophy* 13 (4):505-527.
- BonJour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- Davidson, D (1987). Knowing One's Own Mind. *Proceedings and Addresses of the American Philosophical Association*, 61: 441–58.
- Evans, I., and Smith, N. D. (2012). *Knowledge*. Malden, MA and Cambridge: Polity Press.
- Ghijsen, H. (forthcoming). *The Epistemic Puzzle of Perception: Conscious Experience, Higher-Order Beliefs and Reliable Processes*. Dordrecht: Springer.
- Goldberg, S. (2008). Testimonial Knowledge in Early Childhood, Revisited. *Philosophy and Phenomenological Research* 76(1):1-36.
- Goldman, A. I. (1979). What Is Justified Belief? G. Pappas (ed.), *Justification and Knowledge*, Dordrecht: Reidel.
- Graham, P. J. (2006). Liberal Fundamentalism and its Rivals. *The Epistemology of Testimony*, J. Lackey and E. Sosa, eds., Oxford University Press: 93-115.
- Graham, P. J. (2010). Testimonial Entitlement and the Function of Comprehension. *Social Epistemology*, D. Pritchard, A. Millar, A. Haddock, eds., Oxford University Press: 148-174.
- Graham, P. J. (2012). Epistemic entitlement. *Nous*, 46(3): 449-482.
- Graham, P. (2014). Functions, warrant, history. Fairweather, A. and Flanagan, O., eds., *Naturalizing Epistemic Virtue*. Cambridge University Press: 15-35.

- Lehrer, K. and Cohen, S. (1983). Justification, Truth, and Coherence. *Synthese*, 55(2).
- Lyons, J. (2009). *Perception and Basic Beliefs: Zombies, modules and the problem of the external world*. USA: Oxford University Press.
- Michaelian, K. (2010). In Defence of Gullibility: The Epistemology of Testimony and the Psychology of Deception Detection. *Synthese*, 176 (3):399-427.
- Millikan, R. (1984a). *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press.
- Millikan, R. (1984b). Naturalist Reflections on Knowledge. *Pacific Philosophical Quarterly* 65: 315-34.
- Papineau, David, 1993. *Philosophical Naturalism*, Oxford: Basil Blackwell.
- Pietroski, P. (1992). Intentional and Teleological Error. *Pacific Philosophical Quarterly*, 73: 267–81.
- Proust, J. (2013). *The Philosophy of Metacognition: Mental agency and self-awareness*. Oxford University Press.
- Smith, J. D. (2009). The study of animal metacognition. *Trends in cognitive sciences*, 13(9):389-396.
- Sosa, E. (1993). Proper Functionalism and Virtue Epistemology. *Nous*, 27:51-65.
- Wright, L. (1976). *Teleological Explanations: An Etiological Analysis of Goals and Functions*. London: University of California Press.