**Agent Functionalism**

**DRAFT**

Christoph Kelp

COGITO Epistemology Research Centre

[christoph.kelp@glasgow.ac.uk](mailto:christoph.kelp@glasgow.ac.uk)

1. **Introduction**

Virtue epistemology (e.g. Greco, 2010, 2012; Sosa, 2015, 2021; Zagzebski, 1996). and proper functionalism (e.g. Graham, 2012; Millikan, 1984b; Plantinga, 1993; Simion, 2019) have been among the leading approaches in epistemology. One key reason why they have been so popular is that their central epistemological claims are backed by independently attractive general normative frameworks.

In the case of virtue epistemology (virtue epistemology) this is a framework for performance normativity (PN) (e.g. Sosa, 2015, 2021). Champions of virtue epistemology have not only developed this framework and argued that belief is a kind of performance, but they have also identified central epistemic properties with properties of belief countenanced by PN. To take just one example, a popular view in virtue epistemology is that a justified belief is a belief that is produced by a reliable ability to form true beliefs.

In contrast, champions of proper functionalism (proper functionalism) have employed a framework for functionalist normativity (FN) which has been developed primarily in the philosophy of biology (e.g. Godfrey-Smith, 1994; Millikan, 1984a; Neander, 1991). They have argued that our belief-forming processes have epistemic functions and have identified central epistemological properties with properties countenanced by FN. For instance, a popular view in proper functionalism is that a justified belief is a belief that is produced by a belief-forming process that has the function of producing true beliefs reliably.

Although the two views differ in important ways, the resulting accounts tend to look quite similar. This comes out nicely when we compare the two accounts of justified belief as, respectively, belief from reliable ability and belief from belief-forming process that has the function of producing true belief reliable. Accordingly, it will come as little surprise that champions of each view have claimed that the other view really is a version of their preferred view.

The question of the relation between virtue epistemology and proper functionalism will take centre stage in this paper. Crucially, I will approach this question by taking a closer look at the relation between the two normative frameworks that underpin the two views. In particular, I will argue that PN is a species of FN. That’s good news for champions of proper functionalism. Their normative framework is the more fundamental one.

Another question that I will consider is the question of whether we should opt for a traditionalist or a knowledge first construal of, respectively, virtue epistemology and proper functionalism. Roughly, traditionalist versions aim to reductively analyse knowledge. That is to say, they aim to identify a set of non-circular necessary and sufficient conditions, including justified belief.[[1]](#footnote-1) Knowledge first versions, in contrast, not only abandon the ambition of giving a reductive analysis of knowledge but also reverse the traditional direction of analysis. Rather than analysing knowledge in terms of justified belief, they venture to analyse justified belief in terms of knowledge.[[2]](#footnote-2) I will argue that, once we are clear that PN is a species of FN, there is theoretical reason to go knowledge first.

In fact, I will provide reason to think that once we see that PN is a species of FN and once we are clear that we’ll do well to opt for a knowledge first approach, the view we will end up with reduces virtue epistemology to a version of proper functionalism. Crucially, this view is distinctive in that the functions that are relevant to knowledge and justified belief are agent-relative. The last question I will address is whether there is any reason to favour an agent-relative approach, and I will provide some reason for a positive answer. In this way, the view I am developing in this paper can be described a kind of agent functionalism.

1. **Functionalist Normativity**

I would like to start by outlining FN. To see how FN works, it will be useful to start by taking a quick look at what a function is. As a first step, it is worth distinguishing between two different types of function.

First, a design function (henceforth also ‘d-function’) ultimately turns on the intentions of an agent who may express them in a design plan, which specifies what the functional item is supposed to do and how it is supposed to work. Your coffee maker has a design function which is rooted ultimately in the intentions of its creator and is typically expressed in its design plan.

However, not all functions turn on intentions in this way. Another type of function turns on the existence of a feedback loop involving the functional item and a good functional effect, which it produces in a system. I will henceforth refer to this kind of function as ‘etiological functions’ or ‘e-functions’ for short.[[3]](#footnote-3) The heart is an example of an item with an e-function. It is widely recognised that one (if not the) key function of the heart is to pump blood. Why is that? According to the present account, the answer is that pumping blood is good in that it contributes to the proliferation of genes that are responsible for its existence, by keeping us alive long enough to procreate. The fact that heart pump blood contributes to explaining why hearts exist which, in turn, contributes to explaining why they continue to pump blood. In this way, hearts exemplify exactly the kind of feedback loop characteristic of e-functions.

To see that d-functions and e-functions are genuinely different kinds of functions, note that something can have either one without having the other. The heart has the e-function of pumping blood even if it was not designed and so has no d-function at all. The dustbin of failed inventions is filled with things that have d-functions but no e-functions.

Both d-functions and e-functions have normative import, which can be read off the answers to the following two questions:

1. Does the functional item produce the functional effect?
2. Is the functional item functioning properly?

While functional effects are of course generated in different ways depending on whether the function in question is a d-function or an e-function, once the functional effect has been specified, the resulting normative standard works in exactly the same way for both kinds of function. We ask whether the item fulfils its function. The item meets the normative standard if and only if the answer is yes, no matter whether the function is a d-function or an e-function.

Things are slightly different for proper functioning, where there are different answers to the question as to whether an item is functioning properly, depending on whether the function is a d-function or an e-function. In the case of d-functions, the view is quite straightforward. We ask whether the item works in the way it was intended to by the designer, i.e. in the way the design plan specifies. The item meets the normative standard if and only if it does.

It is also easy to see that things have to be different for e-functions. After all, in case of e-functions, there may just not be a designer. And that’s why there is little reason to think that the prospects for unpacking proper functioning in terms of designer intention are anything but bleak.

What then does proper functioning for e-functions amount to? In order to answer this question, we need a bit of conceptual machinery. We have already seen one crucial concept: function fulfilment. The other two are normal functioning and normal conditions. Roughly, normal conditions are the conditions that obtain in the feedback loop in which the functional item produces the functional effect. And, again roughly, normal functioning is the way of functioning that produces the functional effect in the feedback loop, under normal conditions. To make these ideas a little more concrete, consider the heart once more. Here normal conditions include being hooked up to the arteries and veins of a certain kind of organism in a certain way and normal functioning is beating at a certain rate. Functioning normally (beating at a certain rate) under normal conditions (whilst hooked up to arteries and veins) the heart produces its functional effect (pumping blood).

With these points in play, we can now see what proper functioning amounts to for items with e-functions. In a nutshell, the idea is that proper functioning is normal functioning. When the heart is beating at a certain rate it is functioning properly, it is working as it ought to. In order to figure out whether an e-functional item is functioning properly, then, we need to ask whether it is functioning normally, i.e. whether it is functioning in the way it does when producing the functional effect in the feedback loop, under normal conditions. And, again, the item meets the normative standard if and only if it is.

1. **Performance Normativity**

The key thesis of PN that has been developed most prominently in recent virtue epistemology is that any performance with an aim can be assessed along three dimensions: success, competence and aptness (e.g. Sosa, 2015, 2021). Here is, roughly, what this amounts to. A performance is successful if and only if it *attains its aim*; it is competent if and only if it *produced by the exercise of an ability to attain its aim*; and it is apt if and only if it is *successful because competent*. By way of illustration, consider a free-throw you are taking in game of basketball. According to the present proposal, we can ask whether it is successful, i.e. whether it finds the basket; whether it is competent, i.e. whether it is produced by the exercise of an ability to score baskets; and whether it is apt, i.e. whether it finds the basked because of the exercise of an ability to score baskets.

It is worth noting that success and competence can come apart in the sense that it is possible to produce a performance that is successful but not competent and vice versa. Your free-throw may find the basket even though you are currently just too drunk to produce a competent shot. Likewise, your shot may be competently produced but nonetheless miss the basket, perhaps because of an earthquake that caused the basket to collapse.

Aptness, in contrast, requires both successful and competent performance. It is not possible to produce a performance that is apt but not successful and competent. Crucially, however, aptness requires more than successful and competent performance. Success and competence must be related in the right way. Accordingly, it is possible for a performance to be successful and competent but not apt. To see how, consider your free-throw once more. Suppose that you produce a competent shot. As the ball leaves your hand, it is right on the way to the basket. Unfortunately for you, a gust of wind interferes. As a result, now your free-throw would have missed had it not been for a helper with a wind machine who brings it back on target. Here we have a case in which your free-throw is both successful and competent but not apt. This is because your shot is not successful because competent. It is not that your shot finds the basket because you exercised your ability to make free-throws. Rather, it is because you had a helper with a wind machine in the wings.

1. **The Epistemic Normativity of Belief**

Both normative frameworks have been applied in epistemology. In particular, the idea is to use them to develop novel accounts of knowledge and justified belief.

*Virtue Epistemology*

Virtue epistemologists take our beliefs to be performances that aim at truth. This means that PN can be applied to belief. In particular, we get the result that a belief is successful if and only if it is true, that a belief is competent if and only if it is produced by the exercise of an ability to produce true beliefs, and that a belief is apt if and only if it is true because produced by the exercise of an ability to produce true beliefs.

The next step is to identify normative properties of beliefs as performances that aim at truth with familiar epistemological properties. In particular, according to the one prominent and appealing straightforward proposal, justified belief is identified with competent belief and knowledge with apt belief, alongside the identification of true belief with successful belief already mentioned. One attractive feature of the virtue epistemological approach is that the accounts offered are theoretically very satisfying. True belief, justified belief, and knowledge are but the epistemic species of familiar normative genera. In this way, the view can offer attractive answers to questions about epistemic value. In particular, it can simply derive the value of true belief, justified belief and knowledge from the value of successful, competent and apt performance in general.[[4]](#footnote-4)

In addition, the account predicts that true belief, justified belief and knowledge are related in just the way they are widely believed to be. That is to say, true and justified belief can come apart in the sense that one can have a justified belief that is not true and vice versa. Knowledge entails justified and true belief. And, most importantly, justified true belief does not entail knowledge. In fact, the kinds of cases in which the two come apart have exactly the kind of structure of standard Gettier cases. In this way, rather than attempting to solve the Gettier problem in an ad hoc manner, by devising a condition on knowledge that deals with extant Gettier cases, the view promises to solve the Gettier problem in a theoretically motivated manner. After all, the existence of Gettier cases is already predicted by means of the general normative framework the view employs.

Of course, in order to make these accounts of justified belief and knowledge more precise, we need to say more about the central properties of competent and apt performance. To get clear on what a competent performance is, we need to say more about the nature of ability. The standard view in the literature has it that abilities are agent dispositions to attain the relevant success upon trying in suitably favourable conditions (e.g. Greco, 2010; Sosa, 2021). And to have said disposition is to be such that one would likely enough succeed upon trying, at least in sufficiently normal conditions. Competent performances, then, are performances issuedby dispositions to succeed upon trying.

While there is at least a rough consensus on what an ability is, there is an ongoing debate on what it takes for a success to be *because of* ability. That said, according to the most popular view (e.g. Sosa, 2021), for a success to be because of ability, the success must manifest ability, where this requires not only that one attains the success and performs competently, but also that the conditions for the ability exercised are sufficiently normal. Apt performances, then, are performances that are successful, competent, and conditions for the ability are sufficiently normal.

*Proper Functionalism*

Let’s move on to proper functionalism then. Proper functionalists take our belief-forming processes to have epistemic functions. According to the standard view, this function consists in producing true beliefs, perhaps reliably so (Bergmann, 2006; Graham, 2012; Millikan, 1984b; Plantinga, 1993). Given FN, we get two normative properties for belief. One consists in the fulfilment of the function, the production of a (reliable) true belief. The other depends on what kind of function one takes the epistemic function of belief forming processes to be. If it is a d-function, the relevant normative property corresponds to functioning in the way its creator intended. If, on the other hand, it is an e-function, it corresponds to functioning normally, in the way it does when producing the functional effect in the feedback loop, under normal conditions. Given that normal functioning is proper functioning, the standard proposal now is to identify justified belief with normal functioning.

This leaves the question as to how we should think of the functions relevant to justified belief. Are they d-functions or are they e-functions? I do not mean to answer this question here. Suffice it to say that, if they are construed as d-functions, we’ll have to countenance serious additional restrictions. This is because d-functions come so very cheaply. Anyone might design a device that they intend to produce true beliefs. The mere fact that the device works as intended does not mean that the beliefs it produces are justified. Fortunately, those who want to analyse justified belief in terms of d-functions typically take the designer to be an omniscient, omnipotent and omnibenevolent god. And that will presumably avoid the problem. At the same time, in the case of functions designed by such a god, it is hard to see how they could not also qualify as e-functions. Of course, the e-functions here are set up and perhaps even sustained by the designer. But, for all that, it looks as though they are still going to be e-functions, albeit, perhaps, ones that include a god as part of the feedback loop. In view of these considerations, and in order to keep things as simple as possible, I will simply work on the assumption that the kind of function relevant to justified belief are e-functions.

What about knowledge? Most proper functionalists in the traditionalist camp seem to be less concerned with knowledge. A notable exception is Alvin Plantinga who has defended a proper functionalist account of knowledge for decades. According to Plantinga, knowledge is warranted true belief. This means that the substantive epistemological work we are faced with now lies in the analysis of warrant. Here is his proposal:

A belief B has warrant for S if and only if the relevant segments [of the design plan] (the segments involved in the production of B) are functioning properly in a cognitive environment sufficiently similar to that for which S’s faculties are designed; and the modules of the design plan governing the production of B are (1) aimed at truth, and (2) such that there is a high objective probability that a belief formed in accordance with those modules (in that sort of cognitive environment) is true… (Plantinga 1993, 19)

It's easy enough to see just why the view is proper functionalist. Warrant requires the proper functioning of the processes producing the relevant beliefs. In addition, these processes need to be aimed at truth and they must be reliable. Finally, the agent must be in conditions that are sufficiently similar to the ones for which the process was designed.[[5]](#footnote-5)

Now, for present purposes, it doesn’t matter whether Plantinga gets all the cases right. What I would like to point out is just how similar the account is to the virtue epistemological account of knowledge as apt belief, at least once the latter is precisified along the lines suggested above. In fact, for all I care, the two accounts might as well be identical in extension.

Even so, there is an important difference between the two. On the virtue epistemological account, knowledge is simply an instance of a more general normative property, i.e. apt performance, and the precise account of knowledge drops out of a precise account of apt performance in general. The same is not true of Plantinga’s proper functionalist account of knowledge. The reason for this is that FN features only two general normative properties, to wit function fulfilment and normal/proper functioning. And knowledge cannot be identified with either of these properties.[[6]](#footnote-6) As a result, Plantinga’s proper functionalist account of knowledge differs from the virtue epistemological one in that knowledge does not come out as a species of a familiar normative genus. And this puts Plantinga at a theoretical disadvantage. His view is ad hoc where the virtue epistemological rival is well-motivated. Even if the two views turn out to be identical in extension, there is theoretical reason to prefer the virtue epistemological account.

It is worth noting that there is a way to avoid this comparative disadvantage. Proper functionalists might opt for a distinctively knowledge first version of the view (e.g. Simion, 2019, Forthcoming). According to knowledge first functionalism, the function of our belief forming processes is not to produce true beliefs but to produce knowledge. In the knowledge first incarnation of the view, a belief forming process fulfils its function when it produces a belief that qualifies as knowledge and it is functioning properly when it functions in the way it does when it produces knowledge under normal conditions. It is easy to see that knowledge first functionalism remedies the defect of the traditionalist version of the view: knowledge does come out as a species of a familiar normative genus, to wit, function fulfilment. Note also that, the view analyses justified belief in terms of knowledge. After all, justified belief continues to be identified with belief that is produced by a normally functioning belief forming process. Since knowledge first functionalism unpacks normal functioning in terms of knowledge, we get exactly the kind of reversal of the traditional direction of explanation that is characteristic of knowledge first epistemology.

At this point, it might be thought that the divide between virtue epistemology and proper functionalism boils down to a divide between traditionalists and knowledge firsters. To see why, note that PN features three normative properties, which seems to be exactly what the traditionalist needs. In fact, it may look as though it is at odds with the requirements of the knowledge firster. After all, according to a knowledge first version of virtue epistemology, knowledge is the success condition for belief. While we can of course proceed to analyse competent/justified belief with belief that is produced by an ability to know, it looks as though we now don’t have any familiar epistemic property that corresponds to the property of apt belief. This is somewhat undesirable, especially in view of just how central a role the property of aptness plays in PN. In contrast, FN avoids this difficulty. This is because it features only two normative properties, one of which corresponds to knowledge and the other to justified belief. As knowledge firsters, we do not have an idle normative wheel in our epistemology if we opt for FN.

So, does the decision of whether to go for FN or PN depend on the decision to opt for a traditionalist or a knowledge first approach to epistemology? No. As I have argued elsewhere (e.g. Kelp 2017, 2018) virtue epistemology can accommodate a distinctively knowledge first approach to epistemology. This is because there is independent reason for thinking that there is a class of cases in which a performance is successful if and only if it is apt and that belief is a member of just this class. There is no idle normative wheel in a knowledge first version of virtue epistemology. It’s just that knowledge does double-duty in playing the part of both successful and apt belief.

In sum, it looks as though virtue epistemology at least has a slight theoretical edge over proper functionalism in epistemology, at this stage. After all, virtue epistemology can provide accounts of knowledge and justified belief that are theoretically well-motivated in the sense that they are species of a familiar normative genera no matter whether the view is developed along traditionalist or knowledge first epistemological lines. The reason for this is that PN features three normative properties, which makes it rich enough to do the job. While a distinctively knowledge first epistemological version of proper functionalism can match virtue epistemology on this count, a traditionalist version is less well-placed. This is because FN features only two normative properties, which makes it insufficiently rich for traditionalist needs.

1. **The Functionalist Account of Ability**

Given the arguments from the previous section, it may seem surprising that I aim to defend the thesis that PN is a species of FN. In fact, if the arguments from the last section are correct, this thesis may seem outright absurd. After all, the moral of the story that PN is richer than FN because it countenances a wider range of normative properties. But if that is right, what the proposal seems to amount to is that the richer framework is a species of the sparser one. And that just doesn’t seem to make sense. I completely understand these worries. That said, I’d like to ask you to bear with me for now. I’ll get back to them in due course.

First, I’d like to return to PN. In particular, I’d like to take another look at abilities, which play such a key role in it. Recall that, according to the standard view in the literature, abilities are agent dispositions to succeed upon trying in suitably favourable conditions. I have argued elsewhere in more detail (e.g. Kelp 2017, 2018) that the standard view is too coarse-grained. To see why, note that I may have several ways to attain a success such as making a free-throw in basketball. For instance, I may shoot with my left or my right. It may be that one of these ways qualifies as an ability while the other one doesn’t. I am right-handed and have the ability to make free-throws with my right but not with my left. And, finally, it may be that I am disposed to try in the way that doesn’t qualify as an ability. I may have taken a sacred vow never to shoot with my right again. But I still love playing basketball which is why I am now taking shots with my left. Crucially, I still have the ability to make free-throws, i.e. with my right. However, now that I am only taking shots with my left, I no longer have the agent disposition to succeed upon trying. Thus, the standard view of abilities is too coarse-grained.

The obvious amendment to the standard view at this stage is to construe abilities as relative to ways of performing. What we get then is the idea that abilities are ways of performing that dispose agents to succeed upon using them in sufficiently normal conditions. This will avoid the problem of coarseness of grain.

However, another problem remains. There is a difference between a genuine ability and a mere disposition to succeed. To get this difference into focus contrast your way of producing free-throws, which, suppose, qualifies as an ability to make free-throws, with my way which is to throw basketballs straight up in the air when at the free-throw line, say because I irrationally think that this is the best means to producing a successful shot. It might be that I am in a place in which throwing a basketball straight up in the air reliably disposes me to make free-throws, say because the basket is set up in such a way as to move right under the ball when it is thrown right up in the air. We both have the disposition to succeed when producing free-throws in the relevant ways. Even so, unlike your way of producing free-throws, mine does not qualify as an ability. Abilities are more than ways of producing shots that dispose one to succeed (Kelp, 2017, 2018; Millikan, 2000).

One attractive way of solving the problem is to construe abilities are *agent functions* (henceforth also ‘the functionalist account of abilities’). This is to say that they are ways of producing performances that have the e-function of producing the relevant kind of success for the agent. Of course, e-functions are unpacked as expected in terms of a feedback loop involving the way of producing performances and a relevant success, which it produces for the agent: the way of producing performances explains why the successes are produced and the production of successes explains why the way of producing performances is in place.

On the functionalist account of abilities your way of producing free-throws has the function of making free-throws and so qualifies as an ability. After all, it is part of the right kind of feedback loop involving your way of shooting free-throws and making free-throws: the way of shooting explains why your shots find the basket and the shots’ finding the basket explains why the way of shooting is in place. In contrast, my way of producing free-throws does not have this function. This is because it is not involved in the right kind of feedback loop. In particular, the fact that my shots find the target have no bearing on my way of producing shots. The functionalist account of abilities thus offers an attractive explanation of the difference between genuine abilities and mere dispositions to succeed.

Of course, if the functionalist account of abilities is right, FN applies to abilities. Or, to be more precise, it applies to ways of performing that have the function to produce certain kinds of success. In particular, we can ask whether, on a given occasion, the way of performing fulfilled its function and whether it was functioning properly. For instance, in the basketball case, we can ask whether your way of producing free-throws produced a shot that finds the basket and whether it is functioning normally.

1. **PN as a Species of FN**

What comes to light is that FN is at the very heart of what I take to be the right account of a key property in PN, to wit, abilities. While this is of course promising for anyone who wants to understand PN as a species of FN, a good amount of work remains to be done to attain the ultimate goal. So, let’s return to PN and its three normative properties, success, competence and aptness.

*Competent Performance*

Let’s start with competence. This is the easiest property to make sense of for the functionalist, especially given the above view of abilities as agent functions. Here is my proposal. An agent’s performance is competent, i.e. produced by the exercise of a relevant ability, if and only if it is produced by a normally functioning way of performing that has the function of producing the relevant success for the agent. In other words, competent performance is normal functioning of the agent function that, according to the above account of ability, makes for the ability. Consider, for instance, a competent free-throw, i.e. one that is produced by an exercise an ability to make a free-throws. According to the present proposal, for your free-throw to be competent is for it to be produced by a normally functioning way of shooting that has the function of making free-throws for you, i.e. by a way of shooting that is part of a feedback loop featuring successful free-throws. Note that this proposal is not only appealingly straightforward and plausible, but it also has additional benefit of offering a more detailed account of what it takes for an agent to *exercise* an ability, a point on which extant virtue epistemologies tend to remain quiet. This leaves two properties to account for, success and aptness. I’ll look at success next.

*Successful Performance*

To see how we can accommodate the property of successful performance in FN, let’s first ask how it is that a performance can come to be aimed at a certain success. One obvious answer is by intentional design. This is what typically happens when you take on a free-throw in basketball, for instance. You intend to throw the ball into the hoop. And that’s in part why your performance has the corresponding aim. If you agree that a free-throw has finding the basket as its *constitutive* aim, this is in part why your performance is a token of the type free-throw. For performances that acquire their aim by intentional design, it’s again quite easy to make sense of what it is for a performance to be successful with the resources of FN. Intentional design gives rise to a d-function and a performance with a d-function is successful if and only if it fulfils the d-function. When you intend to throw the basket into the hoop, your performance acquires the d-function of finding the hoop and it will be successful if and only if it fulfils this d-function.

Now, is there any other way in which a performance can come to have a certain aim, one that does not require intentional design? One of the attractive features of FN is that it supports a positive answer to this question. To see how, note first that, for any performance, there is a way of performing. After all, you can’t do something without doing it in a specific way. And if the way that produces the performance has the e-function of producing a certain effect, E, then we can say that the performance itself is successful if and only if it fulfils this e-function, i.e. produces E. By the same token, the performance acquires E as its aim.

Note that very often performances acquire aims in both ways simultaneously. When you try to make a free throw and exercise your ability to do so, that your shot is aimed at finding the target is overdetermined. It acquires this aim in virtue of your intentional design and in virtue of being produced by an ability to make free-throws. At the same time, the two ways can come apart. You may be intending to do something, H, that is so hard that you do not have a way of performing that features producing H among its e-functional effects. If so, your performance will have H as its aim only in virtue of the fact that H is its d-function. On the other hand, a certain way in which a certain kind of bird can fly may have the e-function of energy-efficient migration, say, even though a bird that flies in this way never intended to migrate in an energy-efficient way. If so, the bird’s performance of flying in this way will be aimed at energy-efficient migration, but only in virtue of the e-function of the way of flying at issue here.

Are there other ways in which performances can acquire aims? I am inclined to think that the answer is negative. Given that this is right, however, the result that we get is that we can accommodate the property of successful performance in FN. Here goes. A performance with a certain aim is successful if and only if (i) it has this aim (i.a) as its d-function or (i.b) as its e-function (in virtue of being produced by a way of performing that features this aim among its functional effect) and (ii) it fulfils its function.

That we can accommodate the properties of successful and competent performances within FN may not be all that surprising, especially once we have settled on a functionalist account of the key property of ability. After all, success seems akin to function fulfilment anyway as does competence to normal functioning. That said, it might be thought that aptness is the really hard nut to crack. So, let’s have a go at it.

*Apt Performance*

FN does not feature a normative category corresponding to aptness. If we want to argue that PN is a species of FN, one obvious strategy is inflationary. The key idea here is to augment FN by trying to argue that it does feature a (hitherto overlooked) normative property corresponding to aptness or, at the very least, that it should do so. The obvious candidate for such a property would be something like e-function-fulfilment via normal functioning in normal conditions. Unfortunately, as I will argue in what follows, I think that this would be a mistake. A better strategy is deflationary. The key idea here is to deny that aptness has the kind of special normative significance that champions of PN take it to have. The only properties with special normative significance are success and competence.

To see this, let’s first think about why aptness should have special normative significance in the first place. Here standard cases of human performance are taken to provide key evidence. For instance, in our toy case of attempting to score a free-throw, there does appear to be a genuine normative difference between producing a free-throw that is successful and one that is apt here. It makes sense for this to matter to the person who produces the free-throw. Here is the perhaps easiest way to see this. Suppose you are about to take a shot that’s very important to you, say because it will decide whether or not you make the team. You produce a competent shot, it goes in, and you are on the team. At the same time, you had a helper in the wings who moved the basket to make the ball go in. Without their help your shot would have missed. While your shot is successful and competent it was not apt. Later on, you discover all of this. The key observation about this case is that this matters. It makes sense for us to feel that you have fallen short, and perhaps even that you don’t really deserve to be on the team. These considerations suggest that aptness has normative significance that goes beyond that of success. By the same token, aptness seems to have a legitimate place in PN.

One important reason why it a deflationary strategy is better than an inflationary strategy is that the evidence for the special normative significance of aptness provided by standard cases of human performances can be accounted for with the resources of FN only.

To see how, let’s return to standard human performances such as the basketball case once more. Note that standard cases of human performances are cases of intentional performances. In the basketball case, you intentionally produce a shot that turns out to be successful. Recall that one cannot perform without performing in a certain way. What’s more, when it comes to intentional performances, one cannot perform intentionally without performing intentionally in a certain way. For instance, you cannot intentionally take a free-throw in basketball without intentionally doing so in a certain way (e.g. via an exercise of your ability to do so, with your eyes closed, etc.) Since intentions give rise to design functions what this means is that intentional performances have as their design function not only to bring about a certain success, but also to do so in a certain way. When you intentionally take a free-throw in basketball, you not only design it to be successful, but you also design it to attain this success in a certain way. Typically, in the case of free-throws you design your shot to attain success by throwing it in a way that follows a certain arch-shaped trajectory to the hoop. That said, you could also design it to attain success by throwing it in a way that follows a v-shaped trajectory to the hoop via bouncing it off the ground, etc. What’s important to note is that in order to fully fulfil its design function, an intentional performance must not only attain its ultimate aim, but it must also do so in the way in which it was designed to do so. For instance, for your attempt at a free-throw to fulfil its design-function, it must not only go in, but it must do so in the way that you designed for the shot, by throwing it in a way that follows an arch-shaped trajectory to the basket, say. By the same token, an intentional performance can fail to fulfil its function not only by failing to attain its aim, but also by failing to do so in the way envisaged. And that is exactly what is going on in the basketball case when the helper shows her hand to secure the success. Your performance deviates from its designed path. As a result, even though your shot is successful, thanks to the helper’s intervention, it does not fulfil its design-function. This is because it does not fulfil its function in the way in which you designed it to.

What this means is that FN can make sense of the feeling that, in the case at hand, you have fallen short, even though your shot found the basket. Moreover, this explanation promises to generalise to standard human performances in general, because they are intentional performances. Crucially, there is no need to countenance further normative properties to do so. Rather, function fulfilment is all that’s required. And, of course, that we can explain the apparent normative significance of aptness without invoking a special normative property is an important step towards making the case for a deflationary strategy.

Another important step is that there are cases in which aptness simply does not have the special normative significance that champions of PN would have us think it has. To see this, consider the following case. (Note that it differs from the kind of case that champions of PN use to motivate the normative significance of aptness in that it is not a case of intentional performance.)

While your heart is currently pumping blood through your body, it is at the bottom end of the range for normal functioning. Suppose that this would not normally be a problem. However, you are taking a kind of mediation that causes your blood to flow considerably more slowly through your body than it otherwise would. In fact, pumping at the rate it currently does, your heart won’t support your organism, given the state of your arteries and veins. Luckily you have also been taking a second kind of medication an unexpected side effect of which is that your blood flow more easily than it normally would. In fact, that’s why your heart manages to keep you alive.

In this case, the heart fulfils its function, and it is functioning normally. However, it is not fulfilling its function via normal functioning in normal conditions. But whereas in the cases of your free-throw it matters that your shot falls short of aptness, here it does not matter that your heart falls short of fulfilling its function via normal functioning in normal conditions. It does not make sense for us to feel that your heart’s performance has fallen short, nor that you don’t really deserve to be alive, for that matter. The normative property corresponding to aptness in FN does not have the kind of normative significance that champions of PN would have us think it has.

In conjunction, the above considerations provide good reason for thinking that the deflationary strategy is indeed preferable to the inflationary strategy. After all, what comes to light is not only that the evidence for the special normative significance provided by standard cases of human performance can be explained by FN without invoking a special normative property corresponding to aptness, but also that there is reason to think that aptness simply doesn’t have the kind of normative significance that champions of PN would have us think it has.

In sum, since we have seen that two key normative properties of PN can be recovered by FN, while there is reason not to countenance the third, the path is now clear to understanding PN as a species of FN.

1. **Knowledge First**

Recall that we left our discussion of the epistemic normativity of belief at a point at which virtue epistemology seemed to have an edge over proper functionalism. The reason for this was that PN, with its three normative properties, provided a framework rich enough to accommodate the normative needs, as it were, of both the traditionalist and the knowledge firster. Proper functionalism, in contrast, seemed to cater only to the normative needs of knowledge firsters. This is because it only features two normative properties.

Now, the tables have turned. We have seen that there is reason to think that PN can be understood as a species of FN, that the functionalist’s two normative properties are all that’s needed and that there is reason to doubt that aptness is a bona fide normative property. At this stage, it is only natural to ask what exactly means for the epistemic normativity of belief.

It might be thought that there is at least some reason to go knowledge first. After all, FN features only two normative properties. And that’s just not enough for the purposes of the traditionalist. Insofar as we want key epistemic properties such as knowledge and justified belief to come out as species of broader normative genera, we’ll do well to abandon traditionalism in favour of a distinctively knowledge first approach to epistemology.

While I think that this is basically right, there is a wrinkle that’s worth discussing. Recall that FN can make sense of a normative property that goes beyond aim-attainment, at least when it comes to intentional performances. The reason for this is that in when it comes to intentional performances, one cannot perform intentionally without performing intentionally in a certain way. As a result, in order to fully fulfil its design function, an intentional performance must not only attain its ultimate aim, but it must also do so in the way in which it was designed to do so. This raises the question whether traditionalists may be able to use this normative property to analyse knowledge.

In short, there is excellent reason to think that the answer is no. As a first observation, note that already the virtue epistemological claim that beliefs are performances has faced considerable criticism (e.g. Chrisman, 2012). The further claim that beliefs are intentional performances is clearly too strong. Even if we want to allow that some beliefs are intentional performances, it is hard to deny that at least some beliefs are not intentional performances. The problem is that even beliefs of the latter sort can be gettiered. An animal may have a veridical hallucination, say, in which case they will end up with a belief that is successful and competent but falls short of knowledge. While (traditionalist) PN can explain the absence of knowledge in terms of the absence of aptness, (traditionalist) FN can’t. Since the belief-forming process fulfils its function and is functioning normally, these two properties won’t be able to do the work. And, most importantly, since we are dealing with a case of non-intentional performance, the normative property that goes beyond aim-attainment that FN can draw on to account for intentional performances, can’t do the trick either. On reflection, then, there is reason to think that we’ll do well to abandon traditionalism in favour of a distinctively knowledge first approach, at least provided that we want key epistemic properties such as knowledge and justified belief to come out as species of broader normative genera.

1. **Agent Relativity**

On the view I have been developing, virtue epistemology – or, at the very least, the distinctively knowledge first version of virtue epistemology that I favour – reduces to a kind of functionalism. What is important to keep in mind is that, on this kind of functionalism, the functions that are relevant to knowledge and justified belief are agent-relative. In this way, the view differs from other versions of functionalism. To see this, consider the view that is the perhaps closest competitor to my view and arguably most popular version of knowledge first functionalism in the literature, due to Mona Simion (2019):

A belief is justified if and only if it is generated by a properly functioning cognitive process that has the etiological function of generating knowledge. (2019, 262)

Note that Simion’s version of knowledge first functionalism is agent neutral. What needs to happen for a belief to be justified is that it is produced by a process with a certain function and not, as the agent relative view would have it, by a process that has a certain function for the agent.

The last question I’d like to address in this paper is whether there is any reason to prefer some agent-relative version of functionalism to its agent-neutral rivals. In the remainder of this section, I will provide some evidence for a positive answer.[[7]](#footnote-7)

Let’s start with a familiar kind of case (a variation of Lehrer's (1990) case):

You have recently been exposed to radiation. As a result, you have acquired the unusual power to form spontaneous beliefs about the ambient temperature that are true with a very high degree of reliability. You start forming beliefs about the temperature in this way.

The key observation here is that beliefs you form in this case are not justified and fall short of knowledge. Cases like this one have famously been used to argue against process reliabilist (e.g. Goldman, 1979; Kornblith, 2002; Lyons, 2009) accounts of justification (BonJour, 1980; Lehrer, 1990). Note that one way for process reliabilists to deal with this kind of case is by going knowledge first. While on standard versions of process reliabilism justification has to do with forming beliefs that a reliable in the sense that they produce beliefs with a favourable truth to falsity ratio, on the knowledge first version, what matters would be a favourable knowledge to ignorance ratio. Crucially, while your belief forming process does reliably produce true beliefs, it doesn’t reliably produce knowledge. So, the knowledge first version of process reliabilism can accommodate the intuitive absence of ignorance in this case. Now, another way to go with this case is to opt for a proper functionalist account of justification. After all, while the belief forming process produces true beliefs with a high degree of reliability, it does not have the function of producing true beliefs (reliably) (e.g. Graham, 2012). Accordingly, another way to deal with this kind of case is by abandoning a simple process reliabilism in favour of some kind of proper functionalism.

Why is any of this relevant to the question of whether we should go for an agent-relative or an agent-neutral version of knowledge first functionalism? Given that we are dealing with a view that combines both knowledge first epistemology and proper functionalism, we have a true wealth of theoretical resources to deal with cases like the above. To see the answer to this question, consider the following variation of the case (which is in effect closer to Lehrer’s original case):

I am currently developing an intracranial device that is meant to allow us to keep track of the temperature by producing spontaneous beliefs about the temperature whenever there is a change of at least one degree centigrade. You have agreed to test the product for me and have it installed. Upon trying it out you notice that there are some problems, say, the device is slightly inaccurate. I go back and work a bit more on it, then take it back to you and you give it another test run. Finally, the device works perfectly, and you acquire knowledge about temperature with its help.

Before moving on, note that the device not only produces knowledge about the temperature but also has the e-function of generating knowledge about the temperature. After all, it is part of a feedback loop such that the fact that it is used explains the production of knowledge about the temperature and the fact that it produces knowledge explains why it is around. With these points in play let’s return to the case once more.

Suppose, next, that I take the device out of your skull and implant it in an unwitting subject perhaps during their sleep. When they wake up in the morning, they find themselves forming spontaneous beliefs about the temperature.

Surely, however, these beliefs are no more justified than the beliefs you formed in the earlier version of the case. At the same time, the unwitting subject’s beliefs are produced by a device which has the function of generating knowledge and is currently a normally functioning. Since to function normally is to function properly and since beliefs produced by properly functioning belief forming processes are justified, agent neutral versions of functionalism will struggle to accommodate the absence of justification here.

Note that opting in favour of agent functionalism instead will solve this problem. To see this, note that even if the device in question has the e-function of generating knowledge and even if it has the e-function of generating knowledge *for you*, it does not have the e-function of generating knowledge *for the unwitting subject*. This is because there is no feedback loop of the required sort. In particular, the device does not generate knowledge about the temperature in the unwitting subject. And since for an agent’s belief to be justified it must have been produced by a process that has the function of generating knowledge in the agent, an agent-relative account will predict that while your beliefs about the temperature may well be justified, the unwitting agent’s beliefs aren’t. But, of course, all of this is entirely as it should be. As a result, there is some reason to opt for a distinctively agent-relative version of functionalism.

1. **Conclusion**

This paper has focused on the relation between virtue epistemology and proper functionalism. I have argued that the general framework for the normativity of performances that champions of virtue epistemology have used really is a species of functionalist normativity. In particular, there is reason to believe that the properties of successful and competent performance can be understood in functionalist terms, and there is independent reason to not to countenance aptness as a genuine normative property. With the functionalist construal of PN in play, I have argued that we will do well to opt for a distinctively knowledge first approach. A traditionalist approach to the normativity of belief will need three normative genera if all the key epistemic properties are to be species of broader normative genera. In contrast, a knowledge first approach only requires two. With this point in play, we saw that the resulting knowledge first virtue epistemology boils down to a kind of functionalism. What’s distinctive about this functionalism is that it is agent-relative in that the functions relevant to knowledge and justification are agent functions. While this view differs from at least some versions of view in the literature, we have seen that there is at least some reason to endorse a distinctively agent-relative version of functionalism.[[8]](#footnote-8)

**References**

Bergmann, M. (2006). *Justification Without Awareness*. Oxford University Press.

BonJour, L. (1980). Externalist theories of empirical knowledge. *Midwest Studies in Philosophy*, *5*, 53–73.

Carter, J. A., Gordon, E. C., & Jarvis, B. W. (2017). *Knowledge First: Approaches in Epistemology and Mind*. Oxford: Oxford University Press.

Chrisman, M. (2012). The normative evaluation of belief and the aspectual classification of belief and knowledge attributions. *Journal of Philosophy*, *109*, 588–612.

Christensen, W. D., & Bickhard, M. H. (2002). The Process Dynamics of Normative Function. *The Monist*, *85*(1), 3–28.

Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, *23*, 121–123.

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

Goldman, A. (1979). What is justified belief? In G. Pappas (Ed.), *Justification and Knowledge*. Reidel.

Graham, P. (2012). Epistemic entitlement. *Noûs*, *46*, 449–482.

Greco, J. (2010). *Achieving Knowledge*. Cambridge University Press.

Greco, J. (2012). A (different) virtue epistemology. *Philosophy and Phenomenological Research*, *85*, 1–26.

Kelp, C. (2017). Knowledge First Virtue Epistemology. In A. Carter, E. Gordon, & B. Jarvis (Eds.), *Knowledge First: Approaches in Epistemology and Mind*. Oxford University Press.

Kelp, C. (2018). *Good Thinking: A Knowledge First Virtue Epistemology*. London, UK: Routledge.

Kornblith, H. (2002). *Knowledge and Its Place in Nature*. Clarendon Press.

Lehrer, K. (1990). *Theory of Knowledge*. Routledge.

Logins, A., & Vollet, J. H. (n.d.). *Putting Knowledge to Work: New Directions for Knowledge-First Epistemology*. Oxford: Oxford University Press.

Lycan, W. (2006). On the Gettier problem problem. In S. Hetherington (Ed.), *Epistemology Futures*. Clarendon Press.

Lyons, J. (2009). *Perception and Basic Beliefs. Zombies, Modules, and the Problem of the External World*. Oxford University Press.

McLaughlin, P. (2000). *What Functions Explain: Functional Explanation and Self-Reproducing Systems* (Issue 444). Cambridge University Press.

Millikan, R. (1984a). *Language, Thought, and Other Biological Categories*. MIT Press.

Millikan, R. (1984b). Naturalist reflections on knowledge. *Pacific Philosophical Quarterly*, *65*, 315–334.

Millikan, R. (2000). *On Clear and Confused Ideas*. Cambridge University Press.

Neander, K. (1991). Functions as selected effects: The conceptual analyst’s defence. *Philosophy of Science*, *58*, 168–184.

Plantinga, A. (1993). *Warrant and Proper Function*. Oxford University Press.

Shope, R. (1983). *The Analysis of Knowing: A Decade of Research*. Princeton University Press.

Simion, M. (2019). Knowledge-First Functionalism. *Philosophical Issues*, *29*(1), 254–267. https://doi.org/10.1111/phis.12152

Simion, M. (Forthcoming). Resistance to Evidence and the Duty to Believe. *Philosophy and Phenomenological Research*.

Slaght, R. L. (1977). Is justified true belief knowledge?: A selective critical survey of recent work. *Philosophy Research Archives*, *3*, 1–135.

Sosa, E. (2007). *A Virtue Epistemology. Apt Belief and Reflective Knowledge* (Vol. 1). Oxford University Press.

Sosa, E. (2011). *Knowing Full-Well*. Princeton University Press.

Sosa, E. (2015). *Judgment and Agency*. Oxford University Press.

Sosa, E. (2021). *Epistemic Explanations: A Theory of Telic Normativity, and What It Explains*. Oxford: Oxford University Press.

Williamson, T. (2000). *Knowledge and Its Limits*. Oxford University Press.

Zagzebski, L. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge University Press.

1. One of the key tasks for traditionalist approaches is to provide a solution to the Gettier problem (Gettier, 1963). For surveys of the recent literature on this see e.g. (Lycan, 2006; Shope, 1983; Slaght, 1977). [↑](#footnote-ref-1)
2. The most famous defender of knowledge first epistemology is Timothy Williamson (Williamson, 2000). For more on knowledge first epistemology see e.g. (Carter et al., 2017; Logins & Vollet, Forthcoming). [↑](#footnote-ref-2)
3. Godfrey-Smith, Millikan, and Neander all defend versions of the etiological account of functions. A promising alternative that is the organisational theory of functions which has been defended in e.g. (Christensen & Bickhard 2002; McLaughlin 2000). For present purposes, the question as to which of the two views is the correct one can remain open. [↑](#footnote-ref-3)
4. For more on virtue epistemological approaches to the value of knowledge see e.g. (Greco, 2010; Sosa, 2007, 2011). [↑](#footnote-ref-4)
5. It may be worth noting Plantinga has since acknowledged that even this proposal need refinement; see (Boyce and Plantiga 2012) for more. [↑](#footnote-ref-5)
6. Or, to be more precise, they cannot be so identified so long as we hold on to a traditionalist version of the view, according to which the function is analysed in terms of true belief. More on this below. [↑](#footnote-ref-6)
7. I do not mean to suggest that these considerations settle the case. Simion has adduced reason to favour the agent neutral version. [↑](#footnote-ref-7)
8. The research for this book is supported by a grant from the European Research Council under the European Union’s Horizon 2020 research and innovation programme (grant agreement n° 948356). [↑](#footnote-ref-8)