Knowledge Comes First

Mona Simion

# For *Contemporary Debates in Epistemology*, 3rd edition, eds. Ernest Sosa, Matthias Steup, John Turri, and Blake Roeber, Wiley-Blackwell

**1. Introduction.**

According to the view I will defend here, epistemic justification turns on knowledge. I argue that generating knowledge is the function of our belief formation systems, in virtue of the fact that they are representational systems, and that successful representation via belief is knowledge. Knowledge functions, in turn, give rise to knowledge norms. Since justification turns on norm compliance, epistemic justification turns on knowledge. The bad news is that we are not justified to believe the deliverances of *mere* belief-formation systems: we are only justified to believe the deliverances of *knowledge*-generating systems. The good news is that our belief formation systems *are* knowledge generating systems, so, when they are properly functioning, we are justified to believe their deliverances. Or so I will argue.

In what follows, I will purport to offer three arguments for the three key claims above: that knowledge is the function of our belief formation systems (Section #2), that epistemic justification turns on knowledge (Section #3), and that we are justified to believe the deliverances of our properly functioning belief formation systems (Section #4).

**2. Generating Knowledge Is the Function of our Belief Formation Systems**

I first argue that, since our belief formation systems are representational systems, and since to generate non-knowledgeable beliefs is to fail to successfully represent, the function of our belief formation systems is generating knowledge. Here is, the argument unpacked:[[1]](#footnote-1)

(1) Representational systems have representational main functions.

(2) The representational function of a system is to successfully represent.

(3) If a subject S has a belief formation system, then it is a representational system.

(4) If S has a belief formation system, then its representational main function is to successfully represent (from 1-3).

(5) Belief formation systems successfully represent if and only if they generate knowledge.

(6) If S has a belief formation system, its representational main function is to generate knowledge (from 4 and 5).

I take (1) and (2) to not be in need of a lot of defence. It is eminently plausible that some systems wear their main functions up their sleeves, as it were, in that their dubbing is function-driven. Plausibly, toasters have toasting main functions. Also plausibly, the toasting function of any system is to toast successfully. Washing machines have washing main functions, and the washing function of any system is to wash successfully. Representational systems have representational main functions, and their representational function is to successfully represent.

(3) assumes a view on the nature of our belief formation systems whereby their being representational systems is an essential feature of belief formation systems *qua* belief formation systems. I take this not to be in need of much defence either: after all, beliefs are representational devices, with mind to world direction of fit. That is not to say, of course, that one can know on *a priori* grounds that my eyes etc. generate representations: surely, there could be a world populated by creatures that have the exact same mechanisms that we use for representing, only they use them, say, for digesting. It is an empirical question whether some given system is a belief formation system (Burge 2013). But if a psychological system *is* a belief formation system, its main function is to form representational devices (i.e. beliefs), hence to represent.

(4) follows from 1, 2, and 3.

(5) will likely be the source of much disagreement: representation, some will think, seems to be about truth, not knowledge. Here are, though, a few reasons to think that’s wrong. First, note that a belief that is not knowledgeable (false, or true by luck, or based on a hunch, or Gettierized) is intuitively defective *qua* belief. To see this, note that beliefs that fall short of knowledge in the various ways described above are intuitively criticizable (which, importantly, is not to say that the believers in question are: to the contrary, very plausibly, blamelessly non-knowledgeable believers are not criticisable). In turn, *ceteris absentibus* (i.e., unless there are other reasons for criticism present, that do not pertain to Xs goodness qua X), criticisability indicates lack of attributive goodness: *ceteris absentibus*, it cannot be that something X is criticisable although there is nothing attributively defective with X – i.e., although X is good qua X.

As Williamson puts it, mere belief is botched knowledge (2000).[[2]](#footnote-2) Crucially, note also that knowledge is enough for non-defective belief: *ceteris absentibus* (e.g. in the absence of high practical stakes, or moral considerations against etc.*)*, there’s nothing wrong – qua belief - with a knowledgeable belief that falls short of certainty, for instance. If all this is so, however, (5) follows: since it is plausible that a belief formation system is only successful if it generates attributively good beliefs – beliefs that are good qua beliefs - , and since good beliefs are knowledgeable beliefs, it will follow that our belief formation systems are only successful if they generate knowledge.

To see the plausibility of this further, recall Putnam’s famous ant case:

An ant is crawling on a patch of sand. As it crawls, it traces a line in the sand. By pure chance the line that it traces curves and recrosses itself in such a way that it ends up looking like a recognizable caricature of Winston Churchill. Has the ant traced a picture of Winston Churchill, a picture that depicts Churchill? Most people would say, on a little reflection, that it has not. The ant, after all, has never seen Churchill, or even a picture of Churchill, and it had no intention of depicting Churchill. It simply traced a line […] (Putnam 1981).

This case lends support to content externalism: the ant does not have the concept of Churchill, since she’s never seen or heard of Churchill, and, as a result, the line it draws in the sand does not stand in the right relation to Churchill to count as representing Churchill.

Now, consider the following variation on this case:

Mary has been introduced to Winston Churchill and knows him well. Now, Mary is dancing salsa at a beach party and, unbeknownst to her, by pure chance, as she dances, she traces a line in the sand that ends up looking like a recognizable caricature of Winston Churchill.

Has Mary drawn a picture that successfully represents Winston Churchill? Again, most people would say ‘no’: the line in the sand does not stand in the right relation to Churchill to count as representing Churchill. There is no difference between the ant’s line in the sand and Mary’s: Mary has the concept of Churchill, but the right relation for representation to be instantiated still fails to obtain: the causal chain, as it were, is interrupted on the segment from Mary’s concept to the line in the sand: what explains the correlation is intervening luck (Pritchard 2005): luck that intervenes between Mary’s concept of Churchill and the line she draws in the sand. What this suggests is that representation, just like knowledge, is incompatible, in an important sense, with intervening luck. This, in turn, lends support to (5) above: systems generating luckily true beliefs are not successful representational systems, in virtue of not putting us in cognitive contact proper with the world. Forming lucky true beliefs does not amount to representing the world, just like Mary’s dancing does not count as a representational process, no matter how many times she luckily ends up drawing Churchill-looking lines in the sand. What safeguards against intervening luck is knowledge: Representational belief formation systems are knowledgeable belief formation systems: they are systems that are only successful when they generate knowledge. [[3]](#footnote-3)

**3. Epistemic Justification Turns on Knowledge**

The previous section argued that, since our belief formation systems are representational systems, and since to generate non-knowledgeable beliefs is to fail to successfully represent, the main function of our belief formation systems is generating knowledge. In this section I will argue that if generating knowledge is the main function of our belief formation systems, epistemic entitlement turns on knowledge.

To see that this is so, it’s important, first, to note that functions generate functional norms: functional items ought to work in ways that are, in normal conditions, conducive to function fulfilment. In virtue of their main biological function being to pump blood in our circulatory system, hearts are biologically normatively constrained to work in a way that is, in normal conditions, conducive to pumping blood in our circulatory system – i.e. beating at a particular rate; to see this, note that hearts that fall short are malfunctioning hearts. Since knowledge generating is the main representational function of our belief formation systems, my belief formation systems (representationally) ought to work in a manner that is, in normal conditions, conducive to generating knowledge. Belief formation systems that fail to do so are malfunctioning.

The next step is to notice that justification turns on norm fulfilment: for all *phi* and norms of a particular type T, *phi* is T-justified if and only if *phi* is T-permissible. If so, epistemic justification of beliefs turns on their epistemic normative credentials: one’s belief that *p* is epistemically justified if and only if it is epistemically permissibly formed. Finally, note that, very plausibly, representational norms are epistemic norms. If all this is so – i.e., if representational norms governing our belief formation turn on knowledge, and epistemic justification turns on representational norm fulfilment – it follows that epistemic justification turns on knowledge: beliefs are epistemically justified iff formed in a manner that is, in normal conditions, conducive to generating knowledge.

With the full defence in play, here is the argument unpacked.

(1\*)/(6) If S has a belief formation system, its main representational function is to generate knowledge.

(2\*) If *x*’s function of type T is to *phi*, then *x* (T)-should work in a way that is conducive to *phi-ing* in normal conditions.

(3\*) If S has a belief formation system, then it (representationally-) should work in a way that is conducive to generating knowledge in normal conditions. (from 1 and 2)

(4\*) Representational norms governing belief formation are epistemic norms.

(5\*) One’s belief that p is epistemically justified iff it is epistemically permissibly formed.

(6\*) One’s belief that p is epistemically justified iff it is formed in a way that is conducive to generating knowledge in normal conditions (from 3, 4, and 5).

I have defended (1\*)/(6) in the previous section at length, and I take (2\*) to be in no need of defence; functions are widely taken to generate corresponding norms.

(3\*) follows from (1\*) and (2\*), and (4\*) is eminently plausible. (5\*) is the incarnation for epistemic normativity of belief of the widely accepted deontic thesis for justification: justified *phi*-ing is permissible *phi*-ing. Finally, (6\*) follows from (3\*), (4\*), and (5\*): since belief’s representational norms turn on knowledge, and since belief’s epistemic justification turns on representational norms, the epistemic justification of belief turns on knowledge.

**3. We are Epistemically Justified to Believe the Deliverances of Our Properly Functioning Belief Formation Systems**

This section is the most ambitious so far. It purports to show that we are, *as a matter of fact*, justified to believe the deliverances of our properly functioning belief formation systems. Here it goes:

The first step is to notice that there are two broad ways in which an item – be it trait, artefact, practice etc. – can acquire a function: by design, or by antecedent success. Artifacts paradigmatically have design functions: they are bestowed upon them by the intentions of the designer. Knives have the design function to cut, and planes have the design function to fly. Compatibly with having these design functions, however, artifacts may never achieve success: a particular knife may never successfully cut, a particular plane may never successfully fly etc. Furthermore, an entire artefact kind may display full or systematic failure in design function fulfilment: for the longest time, planes designed to fly failed to fly, for instance. The Museum of Failed inventions stands solid proof to this simple principle: design function need not imply success.

Things look differently when it comes to function acquisition in items that don’t have design functions – like our hearts and cognitive systems. There are two broad theories about how functions get acquired in these cases: according to Robert Cummins, traits’ functions are to be analysed in terms of the contributions that the traits in question make to the sophisticated capacities of the broader system. For example, the heart of a bilaterian animal pumps blood, which in this way contributes to the capacity of the organism to deliver oxygen and nutrients to its tissues. In this way, the heart acquires the function of pumping blood. In contrast, etiological theories of function impose the further requirement on function acquisition, that the contribution in question, in turn, contribute to the explanation of the continuous existence of the trait: the function of the heart, on this view, is to pump blood, because it did so in the past, which was beneficial for the organism, which, in turn, kept the heart alive.

For our purposes here, the differences between these views will be largely inconsequential:[[4]](#footnote-4) what is crucial to note is that both Cummins functions and etiological functions imply past success: in the absence of design, traits of type T cannot acquire a function F without tokens of type T having F-ed (either in the past, or currently F-ing).

If so – and setting aside the option of supernatural design[[5]](#footnote-5) - our belief formation systems cannot have acquired the function of generating knowledge without them having done so successfully in the past.

If I have a perceptual system, then, it has generated knowledge at least once. Now, note that, crucially, systems that have generated knowledge at least once deliver epistemic justification when properly functioning. If this is right, I conclude that, if I have a properly functioning perceptual belief formation system, I am justified to believe its deliverances. Here it goes:

(1\*\*)/(6) If S has a belief formation system, its representational main function is to generate knowledge.

(2\*\*) If a system M has a non-design function to generate F, then M has generated F in the past.[[6]](#footnote-6)

(3\*\*) For all human S, the representational main function of S’s belief formation systems is a non-design function.

(4\*\*) For all human S, if S has a belief formation system, then it has generated knowledge in the past (from 1\*\*, 2\*\* and 3\*\*).

(5\*\*) One is *prima facie* justified to believe that p if one’s belief that p is generated by a properly functioning system that has generated knowledge in the past.

(6\*\*) For all human S, if S’s belief that p is generated by a properly functioning belief formation system, then S is *prima facie* justified to believe that p (from 4\*\* and 5\*\*).

The argument above, crucially, only establishes the conditional: *if* we have properly functioning belief formation systems, then we’re justified to believe their outputs. Compatibly, of course, we might not have any: our belief formation systems may be improperly functioning, or it might be that we have no belief formation systems at all: (6\*\*) is compatible with the idea that the systems that we take to be belief forming systems are, as a matter of fact, not belief forming systems.

Let run through the argument step by step: (1\*\*)/(6) was argued for in Section 2, while the current section made the case at length for (1\*\*): non-design functions imply past success. (3\*\*) is the naturalistic assumption that underlies this paper. (4\*\*) drops out of (1\*\*) to (3\*\*).

(5\*\*) will be the controversial premise of this derivation. I don’t have a knock-down argument for (5\*\*), other than that it intuitively does not suffer exceptions: We are, indeed, intuitively *prima facie* justified to believe whatever our properly functioning knowledge-generating processes deliver, even if they have only generated knowledge in us *once*. The stress is on the *prima facie* here: defeat will make all the difference. If testimony has generated knowledge in me in the past, but Mary has been compulsively lying to me ever since we first met, and I know it, I am not justified to believe her. My *prima facie* justification gets defeated. However, absent defeat, my *prima facie* justification stands.[[7]](#footnote-7) For instance, if my perception delivered just of piece of knowledge and then an evil scientist started generating false perceptual beliefs in me (e.g. by exposing me to extraordinarily veridical holograms), I am justified to believe based on perception.

Is having generated knowledge only once really enough? What if the belief formation system in question did so once twenty years ago, and I’ve forgotten all about it? The answer is ‘yes.’ To see this, all that one needs to do is fully spell out the case: what is the belief formation system in question supposed to be? As soon as this is done, I predict, one of two things will happen: either the system in question will not have the credentials to have generated knowledge in the past, or it will intuitively deliver justification. Even if I’ve forgotten all about the fact that e.g. visual perception generates knowledge – say, because I’ve been isolated in a dark room for the past twenty years - I am still justified to believe based on perception – absent defeat. I conjecture that the reason why this is so is because, in order for a system to generate knowledge at all, it needs to be a fairly sophisticated epistemic machine, with an excellent track record to begin with. Unless I have reason against trusting my eyes, years, testimony of others etc., I am justified to believe their deliverances, if they have generated knowledge in the past.[[8]](#footnote-8)

Finally, the justification claim follows from (4\*\*) and (5\*\*): we are *prima facie* justified to believe the deliverances of our properly functioning belief formation systems; again, that is, on the assumption that we have them.

**5. Conclusion**

The main function of our belief formation systems is to generate knowledge. In turn, they are properly functioning just in case they work in a way that is normally conducive to generating knowledge. When that happens, the beliefs they generate are justified. Since our belief formation systems need to have generated knowledge in the past – on pain of not being belief formation systems to begin with - , it follows that we are justified to believe the deliverances of our properly functioning belief formation systems.

1. Many thanks to Chris Kelp, Peter Graham, Adam Carter, and Jack Lyons, who helped spot a number of problems with several previous versions of this derivation. [↑](#footnote-ref-1)
2. See also (Kelp 2014, 2021) and (Simion 2020) for arguments that knowledge is the goal of inquiry. [↑](#footnote-ref-2)
3. See also (Dretske 1983). [↑](#footnote-ref-3)
4. But see Simion (2019) for an account of epistemic normativity as sourced in inquiry’s etiological function of generating knowledge, and (Simion 2020) for applications to belief and, respectively, assertion and reasoning. [↑](#footnote-ref-4)
5. It’s important to note that a religious, design-based proper functionalism of the kind preferred by Alvin Plantinga will not deliver the optimistic result promised by naturalistic views of the world: after all, we’ve just seen that design functions need not imply success. If so, against arguments presented throughout the history of epistemology, God’s benevolent design need not imply we know anything that we take ourselves to know. Can’t design-based proper functionalism rely on God’s omnipotence, in conjunction with his benevolence, to get the success result? No, due to the epistemic incarnation of problem of evil: epistemic failures and vices are ubiquitous. [↑](#footnote-ref-5)
6. We may want to allow justification to be present at the first moment of knowledge acquisition; if so, we need to take ‘has generated F’ to include currently generating F. Not much will hinge on this. [↑](#footnote-ref-6)
7. In line with most literature of recent years, I bracket generality problems that may and likely will arise at this stage. See (Lyons 2019) for an excellent proposal. [↑](#footnote-ref-7)
8. Thanks to Adam Carter for pressing me on this. [↑](#footnote-ref-8)