Rescuing a Traditional Argument for Internalism

**Abstract:** Early moderns such as Locke and Descartes thought we could guarantee the justification of our beliefs, even in worlds most hostile to their truth, if only we form those beliefs with sufficient care. That is, they thought it possible for us to be *impeccable* with respect to justification. This principle has traditionally been used to argue for internalism. By placing all of the normatively relevant conditions in our minds, we ensure reflective access to what those norms require of us and so sustain the possibility of impeccability (unlike externalism). However, recent challenges to transparency leave this reasoning vulnerable. In response, I show how impeccability can be sustained without requiring transparency. The account only works if we define internal states as those directly accessible to our rational belief forming systems. I argue that this sort of causal internalism, while somewhat revisionary, preserves traditional motivations for internalism while avoiding problems faced by other varieties. The result is a renewed argument for internalism that simultaneously moves us away from access internalism and towards a species of mentalism.

 Early moderns such as Locke and Descartes thought we could guarantee the justification of our beliefs if only we form those beliefs with sufficient care. Describing such views, Alvin Plantinga writes:

We need give no hostages to fortune when it comes to justification; here our destiny is entirely in our own hands. The fates may conspire to deceive me …. Even so, I can still do my epistemic duty; I can still do my best; I can still be above reproach.[[1]](#footnote-1)

The idea is that we can ensure justification not only in the actual world, which we can suppose to be friendly, but even in those worlds most hostile to the truth of our beliefs (e.g., evil demon worlds, brain-in-a-vat worlds, and the like). At best, such guarantees extend only to those beliefs that we are allowed to form with sufficient care. Thus, we should limit our focus to *non-exempt beliefs*, which are those that result from the full and unobstructed operation of our rational belief forming dispositions and so represent our settled takes on the matter. I will refer to this traditional idea—that it is within our power to guarantee the justification of our non-exempt beliefs—by saying, “it is possible for us to be *impeccable* with respect to justification.” We can define impeccability more precisely as:

**Impeccability** – S is impeccable with respect to justification if and only if S possesses a total set of rational belief forming dispositions such that, in all possible worlds where S has that set of dispositions, all of S’s non-exempt beliefs are justified.

What can be said for impeccability? For one, the possibility of impeccability is a feature of justification as originally conceived, since the contemporary notion of justification seems to emerge from early modern thinkers such as Locke and Descartes. For another, the possibility of impeccability makes plausible a close connection between justification and blamelessness—another feature of justification as originally conceived. For if I can follow the norms of justification simply by possessing the proper set of rational dispositions, then, for all non-exempt beliefs, the failure to follow those norms can always be traced back to some deficiency within my intellectual character; and, on some accounts, epistemic blame *just is* the recognition of deficiency within one’s intellectual character, coupled with negative reactive attitudes and responses. Hence, I am blameworthy for all of my unjustified, non-exempt beliefs. There are no excuses.[[2]](#footnote-2) This would be a notable result since externalists have increasingly argued that justification cannot be so closely connected to blamelessness.[[3]](#footnote-3) Another reason to be interested in the possibility of impeccability is that it seems to underlie common reactions to evil demon scenarios—the assumption being that the justification cannot be stripped from these victims by ill-fortune. Finally, many place an interest in justification precisely because they think of it as something for which impeccability is a genuine possibility. These folk (and I count myself among them) prefer to focus on those things that are within their direct power to ensure, and are content to trust God or fate with the rest.

 One can use the possibility of impeccability to motivate internalism about justification. For it seems uncontroversial that externalism about justification always leaves us vulnerable to the winds of fate. Externalism says that what I ought to believe depends on conditions external to me such as the reliability of the processes leading to my belief; but no amount of careful reasoning will ensure that, if I were a brain-in-vat, I would continue to form only those beliefs formed in a reliable manner. In contrast, internalism is ostensibly better positioned to account for impeccability because it places all of the normative relevant conditions inside the agent. Thus, if it is possible to be impeccable with respect to justification, then internalism about justification must be true. We can summarize this reasoning as follows:

 The Impeccability Argument

1. It is possible to be impeccable with respect to justification.
2. It is possible to be impeccable with respect to justification only if internalism about justification is true.
3. So, internalism about justification is true.

The first premise of the impeccability argument is the most controversial. Regarding the second premise, seemingly everyone agrees that externalism provides no hope for impeccability. The concern, in fact, is that there is no hope for impeccability on internalism either; and if impeccability simply isn’t a possibility for humans, then it seems we have decisive reason not to include it in our concept of justification. Hence, most of the work in defending the impeccability argument lies in showing the possibility of impeccability on internalism.

Traditionally, internalists have made the case by appealing to normative transparency. However, recent challenges to transparency have weakened this case considerably. The purpose of this paper is to give an alternative account of impeccability that makes no appeal to transparency. The implications of such an account would be significant. First, it would restore to good health what is nowadays considered an ailing argument for internalism. Second, it would allow the impeccability argument to motivate a version of internalism centered on something other than reflective access. Internalism has increasingly moved away from access internalism and towards mentalism. This paper continues that trend by showing how traditional motivations for access internalism are also enjoyed by a version of mentalism, without relying on controversial claims of transparency.

 In §1, I discuss the traditional account of impeccability and the objections brought against it. In §2-3, I develop an alternative account of impeccability and show how it works for a certain sort of internalism, but not for any form of externalism. Finally, in §4, I discuss the species of internalism arising out of this account and show how it compares to other forms of internalism.

**§1 The Reflective Account of Impeccability**

 The traditional account of impeccability—which I will call “the reflective account”—appeals first to our prime epistemic position with respect to internalist norms.[[4]](#footnote-4) On this account, the mind is *reflectively transparent* to us, meaning:[[5]](#footnote-5)

Necessarily, we are always in a position to know (or justifiably believe) exactly which mental states we are in.

In addition, the support relations that obtain between those mental states and our belief contents are thought to be reflectively transparent as well, meaning:

Necessarily, we are always in a position to know (or justifiably believe) what our mental states do or do not indicate to be true, and to what extent.

Since, on internalist norms, the normatively relevant conditions are exhausted by which mental states one is in and what those states indicate to be true, it follows that internalist norms are *normatively transparent* to us, meaning:

Necessarily, we are always in a position to know (or justifiably believe) what we should and should not believe according to internalist norms.

On the reflective account, it is because internalist norms are normatively transparent, and the norms of justification are internalist in nature, that being impeccable with respect to justification is a possibility for us.

 The problem is that normative transparency is commonly rejected. Regarding the transparency of the mental, there are intuitive counterexamples,[[6]](#footnote-6) empirical critiques,[[7]](#footnote-7) and philosophical objections—the most notable being Timothy Williamson’s anti-luminosity argument.[[8]](#footnote-8) It is also increasingly common to assert that we do not always have unproblematic access to which belief contents our mental states support and to what extent.[[9]](#footnote-9) There are formidable defenses in light of such concerns;[[10]](#footnote-10) but the proponents of transparency are rapidly thinning. For the sake of this discussion, I will grant such critiques (though I do not thereby endorse them).

 If transparency goes, then so does the reflective account. To complete the case, note that there does not seem to be any way of re-grounding the same kind of reflective explanation on more moderate assumptions. For example, one might appeal to the fact that we are *often* in a position to know what internalist norms require of us; but this more moderate control over our beliefs does not guarantee that our non-exempt beliefs will be justified across all possible worlds. Moreover, we have this same limited control over satisfying externalist norms. Insofar as we are generally aware of the cognitive processes leading to the formation of our beliefs, we are, in that respect, also generally aware of whether we are meeting whatever demands externalist norms make of us. So, this more moderate reflective account does not create sufficient separation between internalist and externalist norms to accommodate an argument for internalism on that basis.

 The growing doubts about normative transparency explain why many, including many internalists, have abandoned the possibility of impeccability about justification altogether. For if justification is not normatively transparent, then there will be worlds in which we are, despite our best efforts, ignorant of what we should and should not believe. But then it seems that we have no way of ensuring in those worlds that we are believing as we should and not as we shouldn’t. Amia Srinivasan writes:

Without privileged access to our own minds, there are no norms that can invariably guide our actions, and no norms that are immune from blameless violation. This will come as bad news to those normative theorists who think that certain central normative notions—e.g. the ethical ought or epistemic justification— should be cashed out in terms of subjects’ mental states precisely in order to generate norms that are action-guiding and immune from blameless violation. ... More generally, once we have accepted that our relationship to our own minds lacks the perfect intimacy promised by Cartesianism, we are, for better or worse, left with the view that the normative realm is suffused with ignorance and bad luck.[[11]](#footnote-11)

I respectfully disagree. In the rest of this paper, I will show how we can account for impeccability entirely apart from transparency.[[12]](#footnote-12)

**§2 The Dispositional Account of Impeccability**

The reflective account of impeccability deploys what John Pollock calls “the intellectualist model” of norm guidance.[[13]](#footnote-13) It envisions that we follow norms by consulting explicitly articulated rules dictating how we ought to believe. But what if there was another way to follow norms—one that did not require transparency? Here I will outline Pollock’s non-intellectualist model of norm guidance and show how it might be worked into a non-reflective account of impeccability.

What is this alternative model of norm guidance? Let us begin with an example from the practical realm:

You may make reference to a driving manual when you are first learning to drive a car, but once you learn how to drive a car you do not look things up in the manual anymore. You do not usually give any explicit thought to what to do—you just do it. This does not mean, however, that your behavior is no longer being guided by those norms you learned when you first learned to drive. … The point here is that *norms can govern your behavior without your having to think about them*.[[14]](#footnote-14)

In driving, we see a form of non-intellectualist norm guidance at work. For example, while driving we see brake lights ahead: we immediately brake. We do not consult the rule, “Brake when you see tail lights turn red,” or form any belief about what we ought to do. Nevertheless, that rule guides our driving. What has happened is that, in learning to drive, we have internalized certain patterns of action. More precisely, we developed a set of rational cognitive dispositions that automatically unfold in behavior that adheres to the norms of driving—dispositions that are triggered not by reflecting on the fact that we have certain mental states but by the mere presence of those mental states themselves.

 This set of driving dispositions are not mere reflexes, mind you. With reflexes or other non-rational dispositions, the trigger serves only as the cause of the resulting action. With a rational disposition, however, the trigger is both a cause *and a reason* for the resulting action. In the above example, when seeing tail lights ahead causes us to brake, that experience is the reason or rational basis for our action as well as its cause. This ensures that our driving behavior is not merely in *conformance* with the norms of driving—i.e., we happen to perform the right action in the right conditions—but in *compliance* with it—i.e., we perform the right action *precisely because* we are in the right conditions. When we comply with a norm, we are following or being guided by that norm, even if we do not reflectively consult it. And that is what is happening when we drive. The conditions that dictate we should brake are also the reasons why we do brake, even though we do not reflect on the fact that we have such reasons.

A moment’s reflection on our cognitive processing reveals that we usually form beliefs in the same non-intellectualist way we drive. We are responsive to various mental states as they come into existence, adjusting our beliefs and levels of confidence in accordance with epistemic norms, but rarely do we explicitly consult those norms or what they require of us in light of those mental states. We usually do not reflect on the fact that we have such states at all. For example, while driving we have a perceptual appearance of tail lights: we immediately believe that the car in front of us is braking. We do not consult any epistemic principle to tell us whether that appearance provides adequate support for our belief. We don’t form any beliefs about our perceptual appearance either—any beliefs we have are about the *car* not our *experience* of the car. What’s going on is that the appearance itself *directly guides* our cognition, apart from any reflective beliefs about its existence or normative significance.[[15]](#footnote-15) After all, *the existence* of a mental state or property is a different state of affairs than *reflectively believing* in the existence of a mental state or property, and the rational dispositions resulting in belief formation can be directly triggered by the former as well as the latter. This is a feature of human cognition that Williamson calls “unreflective causal sensitivity to the evidence”.[[16]](#footnote-16)

Importantly, what we are describing here are *rational* belief forming dispositions. This means that our experiences don’t merely cause our beliefs but also serve as the reasonsfor them. What makes the difference between a reason and a mere cause? It seems to have something to do with the *way* in which these appearances cause our beliefs. I will say more about what this way is later on, but for now we needn’t fixate on whythese appearances qualify as reasons, only that they do—and that this remains the case even when we are not reflectively aware that we are having such appearances.[[17]](#footnote-17)

On this non-intellectualist model of belief formation, we can ensure impeccability with respect to our epistemic norms simply by virtue of possessing the right “internalized pattern of behavior that we automatically follow in reasoning”.[[18]](#footnote-18) As an example, consider an epistemic norm N which says of a certain state of affairs X: *Believe p if and only if X obtains*. For S to be impeccable with respect to N, S must meet the following condition:

**Conforming to N** –S possesses a total set of rational belief forming dispositions such that, in all possible worlds where S has that set of dispositions, S forms a non-exempt belief that p if and only if X obtains.

The conformance condition ensures that S forms the non-exempt belief that p if and only if belief that p is called for by N, but this is still compatible with S basing his attitudes on all the wrong reasons. Thus, impeccability also requires that S’s beliefs are *rationally based* on the conditions that permit them:[[19]](#footnote-19)

**Complying with N** – S conforms to N and does so precisely because X does or does not obtain.

If both the conformance and compliance conditions are satisfied, then S is impeccable with respect to N. The crucial point is that both of these conditions can be satisfied without S’s being reflectively aware of whether X obtains or not. S just needs to possess the right set of rational dispositions which will automatically unfold in the right ways in response all and only to the right states of affairs. We will call this “the dispositional account of impeccability.”

 The dispositional account only works if the states of affairs identified as normatively relevant are, in Pollock’s terminology, “directly accessible” to our rational belief-forming systems.[[20]](#footnote-20) Intuitively, directly accessible states are the ones capable of *immediately* or *directly* bringing about belief without requiring that we reflect on our having them. In the driving example, for instance, the appearance of the tail lights brings about our belief without our having to reflect on the fact that this appearance exists. Crucially, then, the “direct accessibility” of these states does not refer to our ability to become immediately aware of them on reflection; it refers instead to the ability of these states to immediately cause us to believe. Thus, the first and most important feature of directly accessible states is that they can trigger our rational belief-forming dispositions merely by existing. Another way to put this is:

**Direct Accessibility (first pass)** – X is directly accessible to S’s rational belief-forming systems only if X (rather than reflective awareness of X) can be the proximate cause of one of S’s rationally formed beliefs.

 Since we are dealing with *rational* belief-forming dispositions—ones whose causes are also reasons for belief—we are operating here on what Wedgwood calls “the folk-psychological level of explanation”. Wedgwood explains that the folk-psychological order of explanation operates at “the personal, mental level,” and so features mental states “of the sort that are referred to in everyday folk-psychological discourse,” such as beliefs, perceptions, appearances, and the like. These kinds of explanations do not appeal to “subpersonal mechanisms or modules in the brain or anything of that sort”.[[21]](#footnote-21) There is, of course, a different explanatory order within which the proximate causes of our beliefs will be brain states; but the states that interest us here are the proximate causes of belief within folk-psychologicalexplanations, since this is the level at which beliefs are evaluated with respect to reasons and rationality. Hence, we must add to our understanding of direct accessibility as follows:

**Direct Accessibility (second pass)** – X is directly accessible to S’s rational belief-forming systems only if X can be the proximate cause of one of S’s rationally formed beliefs within the folk-psychological explanation of that belief.

 Thirdly, some explanations within the folk-psychological order are articulated in more detail than others. Some states of affairs will be described as the proximate causes of our beliefs in less-than-fully-articulated folk-psychological explanations, but will not show up as proximate causes within fully-articulated explanations. For example, we might say truly that, “The man believed his laundry was done because the dryer went off.” However, we can add more detail to this explanation by further specifying, “The man believed his laundry was done because *he heard* the dryer go off.” Thus, while the dryer going off might be a proximate cause of the man’s belief in some sense, it is not the *most* proximate cause of his belief—there are intermediary causal states by virtue of which the dryer’s going off caused the belief. We can confirm this by noting that when hearing the dryer go off comes apart from the dyer in fact going off, belief tracks with the former. For instance, if the dryer goes off but the man doesn’t hear it, then he does not form the relevant belief. When it comes to directly accessible states, the ones we are looking for are the *most proximate causes* of our beliefs. These will be the ones that feature as proximate causes within *fully-articulated* folk-psychological explanations, meaning that they include “all the intervening steps that can be captured at the folk-psychological level”.[[22]](#footnote-22) Thus, we reach the following:

**Direct Accessibility (third pass)** – X is directly accessible to S’s rational belief-forming systems if and only if X can be the proximate cause of one of S’s rationally formed beliefs within the fully-articulated folk-psychological explanation of that belief.

As the dryer example illustrates, the most proximate causes of our beliefs on the folk-psychological level of explanation are always going to be mental states. In other words, “X” will always make reference to the presence or absence of some mental state M.

 Finally, for the dispositional account to work, all of the defining characteristics of the normatively relevant mental states need to be ones which can directly affect belief. Otherwise, a mental state M\* with the same causal properties as the normatively relevant state M could bring about belief even when M is absent, inhibiting perfect conformance. Furthermore, belief would not be brought about when S has M *precisely* *because* S has M (as opposed to M\*), preventing compliance. Hence, the mental states that are directly accessible in the relevant sense will be ones characterized exclusively by those properties by virtue of which they are capable of serving as the most proximate causes of our beliefs. Returning to our earlier example, say the man heard the dryer go off on a Wednesday. We can truly say then that, “The man believed his laundry was done because he heard the dryer go off *on a Wednesday*.” Nevertheless, the mental state *hearing the dryer go off on a Wednesday* is not directly accessible in the relevant sense because that state is characterized by features that make no difference to whether belief is immediately formed. A sure sign of this is that hearing the dryer on a Wednesday makes no difference to whether hearing the dryer immediately brings about belief or not. If the man heard the same sound on a day other than Wednesday, he would have formed the same belief. So, the state *hearing the dryer go off on a Wednesday* is not directly accessible to our rational belief-forming systems.[[23]](#footnote-23) Wedgwood would go so far as to say that this mental state is not part of the correct explanation of the belief at all. He writes:

[I]n any correct explanation there must be a certain sort of proportionality between the explanandum and the explanans. The explanans must be sufficient in the circumstances to produce the explanandum; but it also must not contain any irrelevant elements that could be stripped away without making it any less sufficient to produce the explanandum.[[24]](#footnote-24)

If Wedgwood is correct, then all the better for my argument. In that case, our third pass at direct accessibility would be equivalent to our final characterization below, since none of the mental states featured in our fully-articulated folk-psychological explanations would be characterized in ways that are irrelevant to their ability to bring about belief. Some, however, may argue that states such as *hearing the dryer go off on a Wednesday* do in fact explain why the man believed his laundry was done. These states are simply *partially superfluous* explanations in that they contain more than what is strictly required to explain why the belief came about. To accommodate this sort of position, we can give our full characterization of directly accessible states as follow:

**Direct Accessibility** – X is directly accessible to S’s rational belief-forming systems if and only if X can be the proximate cause of one of S’s rationally formed beliefs within the fully-articulated, entirely non-superfluous folk-psychological explanation of that belief.

 Not all mental states are directly accessible to us in the defined sense. This fact makes it impossible for us to be impeccable with respect to all norms of belief, at least if we are relying on the dispositional account. Return to the norm: *Believe p if and only if S has M*. Humans can guarantee conformance to that norm if and only if (i) M is directly accessible to us, or (ii) M\* is directly accessible to us and, necessarily, M\* obtains if and only if M obtains. Otherwise, it will be possible that M obtain and for our belief forming dispositions not to be triggered, or for those dispositions to be triggered when M does not obtain. In short, the dispositional account can satisfy the conformance condition on impeccability vis-à-vis a norm if and only if the normatively relevant conditions are either themselves directly accessible to us, or else necessarily correlatewithconditions that are. It is even harder to satisfy the compliance condition on impeccability. For compliance requires that our beliefs be rationally based on the states which make them permissible. In this case, that requires believing that p if and only if S has M *and precisely because* S does or does not have M. But this can be guaranteed only if M itself is directly accessible to us. Thus, the dispositional account can preserve impeccability with respect to a norm if and only if the normatively relevant conditions are themselves directly accessible.

**§3 Why this Only Works on Internalism**

In this section, I argue that the dispositional account preserves the possibility of impeccability with respect to justification ifinternalism is true, but not if externalism is true. Along the way, I develop a more specific characterization of internalism which arises organically from our discussion thus far.

*3.1 Impeccability and Internalism*

 To determine whether we can be impeccable with respect to internalist norms, we must first ask what exactly makes for an internalist norm? As a first pass, we said that internalist norms are those whose normatively relevant conditions are all internal to the mindof the subject. That is, the normatively relevant conditions consist exclusively in the presence or absence of certain mental states. This much internalists agree on. But exactly *which* mental states are normatively relevant and why?

*Access internalism* is the original form of internalism as endorsed by traditional figures like Locke and Descartes. It says that the normatively relevant mental states are those to which the subject enjoys a special kind of reflective access—namely, one is always in a position to know (or justifiably believe) whether they exist and perhaps also what they do or do not indicate to be true. I propose that, most fundamentally, what makes something normatively relevant for access internalists is that the subject is able to *take it into account* in belief formation, and the specific manner in which the subject takes it into account is by being reflectively aware of it. Given the challenges to mental transparency cited above, many internalists have moved away from access internalism. Since we are granting those challenges here, we will do the same.

The main alternative to access internalism is *mentalism*. Mentalism says that justification, or how one ought to believe, supervenes on one’s mental states, but denies that we need any special access to those mental states.[[25]](#footnote-25) Complicating matters are the existence of factive mental states like knowing or perceiving—ones characterized in part by their relation to obviously external states of affairs. If these kinds of mental states are included in the supervenience base, then the resulting mentalism would not be internalist in any recognizable sense. Thus, to be an instance of internalism, we must limit the sort of mental states that are included in the supervenience base. Usually, it is restricted to all and only non-factive mental states. Such non-factive states are internal to the subject in a natural enough sense to qualify this as a form of internalism, at least on first blush. Let’s call this form of mentalism “non-factive mentalism”.

One worry about this form of mentalism is that the subject may have no way of registering reflectively or otherwise whether certain non-factive states exist or not. In allowing such states to count as normatively relevant, however, we contravene what I proposed as the fundamental idea behind internalism as originally conceived: that the normatively relevant conditions are those that the subject can take into accountin belief formation. Thus, Michael Bergmann understandably questions whether we ought to consider this a form of internalism at all.[[26]](#footnote-26) Worse still, by allowing some normatively relevant conditions to be completely undetectable (even to our automatic processing systems), non-factive mentalism cannot use the dispositional account to sustain impeccability with respect to justification. That account requires that all normatively relevant conditions be directly accessible to our automatic processing systems. The consequence is that non-factive mentalism cannot claim for itself the motivations for internalism under discussion in this paper.

There is, however, a different form of mentalism inspired by John Pollock that can take advantage of the dispositional account. On this approach, the mental states included in the supervenience base are all and only those that are directly accessible to the subject’s rational belief-forming systems. I will argue momentarily that directly accessible states are a subset of non-factive mental states; thus, if non-factive states are “internal” in a natural sense, directly accessible states are even more so. For while there are non-factive mental states whose presence or absence can go completely undetected by us, and which have no potential for affecting belief formation, this is not true of directly accessible mental states. We invariably detect their presence or absence, and do so in a way that allows us to take them into account in forming beliefs. To be clear, the sense in which we “detect” directly accessible states is not by having actual or potential second-order awareness of them (though we can become aware of them in this way); rather, we detect them in the sense that our rational belief-forming dispositions are sensitive to their presence or absence. This allows us to take them into account in the sense that they wield causal influence over which beliefs we form. Let us define “causal internalism” as a form of mentalism which says that justification (or how we ought to believe) supervenes on all and only one’s directly accessible mental states.

Using the dispositional account, it is easy to see how impeccability can be sustained with respect to justification if causal internalism is true. What we need for impeccability is that all of the normatively relevant conditions be directly accessible to our automatic processing systems. This is obviously satisfied on causal internalism since the supervenience base is specifically restricted to directly accessible mental states.

*3.2 Impeccability and Externalism*

 On first blush, it may seem that the dispositional account also provides externalists with an avenue towards preserving impeccability. Why *can’t* I be disposed to form only reliable beliefs, for instance, across all possible worlds in which I am similarly disposed? The problem is that, on externalism, the normatively relevant conditions are neither themselves directly accessible (as required for compliance) nor do they necessarily correlate with things that are (as required for conformance). But both of these are required for impeccability. Let’s look more closely at each of these two conditions.[[27]](#footnote-27)

 Let’s start with conformance. Consider an externalist norm that says of some external, non-directly accessible state of affairs X: *Believe p if and only if X obtains*. To maintain the possibility of impeccability, it would need to be the case that there is some kind of directly accessible state M such that, necessarily for S, states of kind M exists if and only if X obtains. However, the plain fact is that things on the inside do not perfectly track things on the outside. It seems there will always be possible scenarios in which an instance of M is brought about in the absence of X. As an example, a simple version of process reliabilism says that S’s belief is justified if and only if that belief results from a reliable process.[[28]](#footnote-28) Clearly, there are no directly accessible mental states that always and only lead to reliably-formed beliefs. The same directly accessible state that causes a reliable belief in normal worlds will cause an unreliable belief in evil demon worlds. The problem generalizes to all similar version of externalism because directly accessible states can always be caused in different ways. Skeptics take advantage of this fact in designing skeptical scenarios: they preserve the same kind of directly accessible states that one would have in a normal environment while swapping in abnormal causes. It seems, then, that there are no directly accessible states that necessarily correlate with whatever external conditions the externalist identifies as normatively relevant.[[29]](#footnote-29) Thus, we cannot guarantee even conformance to externalist norms.

 The only kind of externalism that has any chance of satisfying the conformance condition is the sort that Pollock calls “norm externalism”.[[30]](#footnote-30) This sort of view adopts internalist rules but then legitimizes those rules by appealing to the external benefits of following them. For example, a version of normal worlds reliabilism says that a belief is justified for S if and only if it results from a belief-forming process that is reliable *in normal worlds*.[[31]](#footnote-31) Such a theory does not require that the process producing S’s belief *actually be* reliable, but only that this process *would be* reliable in worlds that are “normal” for S. Accordingly, S can ensure justification across all of the relevant worlds (including evil demon worlds) by being disposed so as to believe reliably in normal worlds. This only works, however, if there is a set of rational belief-forming dispositions that prove reliable in normal worlds. The problem is that this needn’t always be true. If S’s belief-forming processes are not reliable even in worlds that are normal for S (say S’s home environment is an evil-demon world or, more mundanely, S’s faculties simply evolved to be unreliable), then S will not be able to guarantee the justification of S’s non-exempt beliefs no matter how carefully S reasons.

 Regardless, even if externalism could somehow secure conformance, the second and more decisive problem is that we could not *comply* with externalist norms—that is, our beliefs would not always be brought about *precisely because* the relevant external conditions obtained. This is because the conditions identified by externalist theories as normatively relevant are not directly accessible to us. Wedgwood effectively argues for this principle by showing that the proximate causes of belief in fully-articulated folk-psychological explanations are always internal facts about a thinker’s mental states,[[32]](#footnote-32) but this argument should be unnecessary for the majority of externalists. On most versions of externalism, it is evident that the conditions to which they are appealing are not directly accessible. Obviously, proponents of those theories will dispute whether it is important that normatively relevant conditions be directly accessible, but they will not pretend that facts about, for instance, the reliability of one’s perceptual processes somehow show up as the proximate causes of our beliefs in fully-articulated folk-psychological explanations.[[33]](#footnote-33) Such externalists do not dispute, for instance, that evil demon victims would believe the same things as their ordinary counterparts, even if the external conditions to which they appeal were absent.[[34]](#footnote-34) Hence, most externalists will concede that their favored external conditions are not directly accessible to our rational belief forming systems.

 The only externalists who might contest our having no direct access to the external are disjunctivists and some in the knowledge-first camp, although even these should, I think, eventually concede the point. The sort of theorists I have in mind are those who endorse norms such as: *Believe p (perceptually) if and only if one perceives p* or *Believe p if and only if one knows that p*. Perception and knowledge are factive mental states, meaning that to perceive or to know p is to stand in a kind of relation that is possible only if p is true. Intuitively, it might seem that states like perception or knowledge are directly accessible to us. After all, we often say things like, “I believe p because I know q” or, to return to our earlier example, “The man believes his laundry his done *because he hears the dryer go off*” (hearing being a species of perception).

 If we look more closely, however, we can see that perception, knowing, and other factive mental states do not meet the conditions on direct accessibility. This is because those states are characterized by features that make no difference to whether belief comes about, and so at best constitute partially superfluous explanations of belief. For example, in the dryer case, it makes no difference whether the man *hears* the dryer go off or whether it merely *appears* to him that the dryer is going off (a state phenomenally indistinguishable from hearing the dryer go off)—he will believe or not believe the same way in both cases. Michael Huemer makes the point using a different example:

[T]he appearance probabilistically screens off my belief from the external fact. That is, given that I experience exactly the sort of appearance I am now experiencing, the probability of my forming the belief that there is a cat is unaffected by the actual existence or non-existence of the cat. If I have this sort of appearance caused by a cat, I will believe that there is a cat; but equally, if I have the same appearances when no relevant cat exists, I will believe that there is a cat.[[35]](#footnote-35)

Even disjunctivists and knowledge-firsters concede this, as shown in their admission that evil demon victims will believe in exactly the same ways that their ordinary counterparts do despite lacking the same factive mental states. It follows that an entirely non-superfluous folk-psychological explanation of belief will not include mention of factive states like perception or knowledge, since some of the defining features of these states play no causal role in directly bringing about belief. To be clear, I am not saying, along with Wedgwood, that perception or knowledge do not feature in fully-articulated folk-psychological explanations of belief at all. I am not even denying that perception and knowledge are sometimes the most proximate causes of our beliefs. I am making the more modest point that if and when such states do serve as the most proximate causes of our beliefs, the properties by virtue of which they do so are not those that qualify them as factive states. Hence, such states are not directly accessible in the manner required to take advantage of the dispositional account.

 To drive the point home, consider the following example from Miriam Schoenfield.[[36]](#footnote-36) The Colonial American lookout Paul Revere makes the following plan:

**The Revere Plan** – Believe that the British are coming by land if he perceives the British coming by land; believe they are coming by sea if he perceives them coming by sea.

Let us imagine, however, that the British had taken to using decoys—objects that could make it appear as though they were coming by land or sea when they were not. The Colonialists are fretting about the reliability of Revere’s plan in light of such decoys when someone interjects: “I don’t see the problem. Revere’s plan is to believe that the British are coming by land if he *perceives* them to be coming by land, not if it only *appears* to him that they are coming by land. And the same for the sea. So, where’s the issue?” The issue, of course, is that Revere’s plan will in fact result in Revere believing that the British are coming by land or sea whenever it appears to him that they are coming that way (and not only when he genuinely perceives this). Indeed, the issue with the interjector’s question is so obvious to us that those listening would likely take it for an ill-timed joke. However, if perception *could be* the proximate cause of belief within a fully-articulated, entirely non-superfluous folk-psychological explanation, then the interjector’s question wouldn’t be silly at all. There would be no principled reason why Revere couldn’t have a rational disposition that was directly triggered by perception but not by a phenomenally indistinguishable appearance. What we observe, however, is that anyone who is disposed to form the belief that p upon perceiving p will also form the belief that p upon apparently perceiving p. Unless we concede that the same kinds of directly accessible states are prompting belief in both cases—in particular, states characterized by what it feels liketo perceive something—then we are utterly at a loss to explain this connection.

 Indeed, it is so evident that factive states such as perception and knowledge are not directly accessible to us that denying this is in itself a strong reason to reject the theory in question. Fortunately, we do not needn’t ascribe such views to proponents of disjunctivism or knowledge-first. Even these can and should accept that we have no direct access to factive states. Timothy Williamson, for example, concedes that we are “causally insensitive to any difference in evidence” between ordinary cases of knowledge and skeptical scenarios.[[37]](#footnote-37) Likewise, Harold Langsam writes of John McDowell:

McDowell, although a disjunctivist about experience, is not only committed to saying that perceptual experiences share a categorical property with illusory experiences (the property of being appearances), but he is also committed to saying that it is this shared property that grounds the power of perceptual experiences to produce beliefs.[[38]](#footnote-38)

Even on McDowell’s view, then, factive states are at best partially superfluous explanations of belief. How are these concessions consistent with knowledge-first and disjunctivism? Because Williamson goes on to insist that “Our causal insensitivity to any difference in evidence between the two cases does not show that there is no difference in evidence between them.”[[39]](#footnote-39) In like manner, Langsam defends McDowell’s epistemology by claiming that the inaccessibility of factive states does not prove their normative irrelevance. He writes:

I deny … that *only* properties that ground the experience’s causal power to produce beliefs can be justification-conferring properties. … [T]here is simply no reason why *every* justification-conferring property of an experience must also be a property that grounds its causal power to produce beliefs.[[40]](#footnote-40)

Now, a central thesis of *this* paper is that the factors relevant for justification *are* limited to the directly accessible (because only then can we be impeccable with respect to justification). Nevertheless, the lesson to take from Langsam and Williamson is that externalists’ real objection is to the *normative* *claim* that only directly accessible states are epistemically relevant vis-à-vis justification, not to the *causal claim* that factive mental states are not directly accessible.[[41]](#footnote-41)

 Because external states of affairs, including the existence of factive mental states, are not directly accessible to us, it follows that we cannot be disposed so as to comply necessarily with externalist norms in the manner required for impeccability. It turns out, then, that the impeccability argument for internalism has teeth after all. If justification is the sort of thing that we can guarantee for our non-exempt beliefs so long as we reason correctly, then internalism about justification must be true. For impeccability is possible only if we limit the normatively relevant conditions to those that are directly accessible to us—namely, to the internal.

**§4 Causal Internalism**

In closing, I want to add some final remarks on causal internalism, mostly to address the worry that it is not a genuine form of internalism at all. My position is that while causal internalism is somewhat revisionary, it remains faithful enough to the core features of internalism as traditionally conceived that the label is not misapplied. The importance of our conclusion does not hinge on this debate, of course. We could easily reframe the discussion as a defense of mentalism about justification with no less interest. Whatever we call it, the requirement that justification supervene on directly accessible mental states would rule out every major theory of justification that marches under the banner of externalism (and some that march under the banner of internalism as well)—and that conclusion is plenty significant on its own.

My position, then, is that causal internalism *is* a form of internalism as that term is generally used. It’s just one which reorients the focus away from reflection and towards causation (specifically, the particular sort of causation involved in rational basing). As Schoenfield aptly puts it, “internalist norms deserve our attention, not because of any special epistemic access that we have to internalist conditions (though perhaps we have that too), but rather because of a special causal role that we should expect internalist conditions to play”.[[42]](#footnote-42) Through this reorientation, I propose that casual internalism elegantly splits the difference between other forms of internalism, preserving their advantages while escaping their disadvantages. Let’s examine how.

Access internalists correctly recognized that for something to be relevant to justification, the subject must be able to *take it into account*. Where they erred is assuming that this must take the form of being reflectively aware of it. However, in retreating from this assumption, non-factive mentalism lost the original insight altogether. Causal internalism, on the other hand, preserves it in a suitably moderated way. On causal internalism, mental states are taken into account by having direct causal influence over rational belief formation. This causal influence, keep in mind, is not merelycausal. When it comes to rational dispositions, the triggering state serves as a reason as well as a cause.[[43]](#footnote-43) And, clearly, if a mental state serves as the rational basis for a belief, then that state is taken into account by the subject in an especially robust sense. Thus, by reconceiving of access in terms of (a special kind of) casual sensitivity rather than reflective awareness, casual internalism captures some of the central motivations for access internalism while avoiding its liabilities.

 Is there any awareness requirement on causal internalism? Bergmann argues that without an awareness requirement, it’s hard to see casual internalism as genuinely internalist. He suggests that on causal internalism:

[It’s possible that] most of our justified beliefs are justified in virtue of internal states we aren’t aware of (or even potentially aware of). Thus, most of our justified beliefs are justified despite the fact that we aren’t aware of even potentially aware of *anything at all* contributing to their justification. To classify such a view as an internalist view does *not* seem to fit at all with the way the term ‘internalism’ has been used in the literature.[[44]](#footnote-44)

It's not entirely clear that conscious awareness should be required for internalism. Even without any awareness requirement, causal internalism rules out *all* common externalist theories of justification and limits normatively relevant conditions to the states within the agent’s own mind that, on the personal level, bring about the agent’s beliefs. So, if forced to pick a label, “internalism” seems more appropriate than “externalism”.[[45]](#footnote-45)

 Nevertheless, let us assume for the sake of argument that an awareness requirement is crucial to internalism. Plausibly, casual internalism can accommodate this. Here’s how. There are reasons to think that rational belief formation (when a belief is formed on the basis of *reasons* rather than brutely caused) is always triggered most immediately by what Smithies calls “phenomenally individuated” mental states.[[46]](#footnote-46) Smithies elaborates:

These “phenomenally individuated” mental states include not only your experiences, which are individuated by their phenomenal character, but also your standing beliefs, desires, and intentions, which are individuated by their dispositions to cause certain kinds of experiences under experiential conditions. At the same time, these exclude your “subdoxastic” mental states, which are individuated by their role in unconscious computational processes, and all your mental states that are externally individuated by their relations to the external world.[[47]](#footnote-47)

Phenomenally individuated states such as these are ones that we are or could be aware of (on the first-order level) simply by virtue of having them. We can sometimes also become *reflectively* aware of them (on the second-order level), but that is something different. Bonjour describes this difference by contrasting the “built-in,” “constitutive awareness” that we have of conscious phenomenal states with “higher-order awareness” of them.[[48]](#footnote-48) Here’s the crux: If phenomenally individuated states are always the proximate causes of rational belief formation within the fully-articulated folk-psychological explanation of our beliefs, then, given casual internalism, we always are or always could beconsciously aware of our reasons for belief—reasons which, good or bad, determine whether our belief is justified. We aren’t always *reflectively* aware of those reasons, but *that* isn’t a requirement for genuine internalism, or at least it shouldn’t be.

 So, are phenomenally individuated states always the most proximate causes of our rationally formed beliefs within folk-psychological explanations? It seems so. Ordinarily, non-basic beliefs are directly caused by other beliefs, and basic beliefs are directly caused by conscious experiences (such as those present in perception, memory, reason, and introspection); and both beliefs and experiences are phenomenally individuated states. What else might serve as the most proximate cause of our beliefs?

 Neuroscientists might identify certain unconscious brain processes as immediately causing belief, but we are looking for the most proximate causes within *folk-psychological* explanations. Such explanations operate on the personal level and so will not include subpersonal brain states. Similarly, Bergmann proposes that something like radio signals might serve as the most proximate cause of our beliefs.[[49]](#footnote-49) The counterexample fails, however, in one of two ways. If the radio signals cause belief by directly rewiring our brains, then there won’t be a folk-psychological explanation of the belief at all. There is no folk-psychological explanation because the belief is not rationally formed. The radio signals *bypass* our rational belief forming dispositions rather than trigger them. Like subpersonal brain states, such radio signals will still count as causing our beliefs, but only within an altogether different order of explanation. On the other hand, if the radio signals do manage to show up in a folk-psychological explanation of the belief, it will only be by causing us to have a conscious mental state (like an experience) that then serves as the most proximate cause of the belief. Simply saying “S was struck by radio signals” does nothing to make intelligible or unsurprising on the personal level why S came to believe p unless something like the following is also true: “S was struck by radio signals *which made it appear to S that p*”.[[50]](#footnote-50) We can generalize the lesson here: Nothing can serve as the most proximate cause of belief within a folk-psychological explanation unless it is operating on the personal level, and nothing operates on the personal level unless it is a conscious mental state with phenomenal properties.

 What’s more, it seems that nothing other than these phenomenal properties is relevant to belief-formation. Evil demon cases prove the point. It goes unchallenged that evil demon victims and their ordinary counterparts will share all and only the same rationally formed beliefs.[[51]](#footnote-51) This is because the victims and their counterparts are stipulated to be *phenomenal duplicates*—what it feels like to be one is exactly the same as what it feels like to be the other—and the unstated assumption is that what we believe supervenes on how things feel to us. We are on solid ground, then, in assuming that the most proximate causes of our beliefs on the folk-psychological level will always be phenomenally individuated mental states. That is, only phenomenally individuated mental states are directly accessible to us.

 There may also be another, in-principlereason why phenomenally individuated states are always the most proximate causes of our rationally formed beliefs. When it comes to rational belief forming dispositions, the triggering states must cause those beliefs in a special sort of way—one that allows the triggering states to serve as *reasons* for the resulting beliefs as well as causes. What could this be? A plausible proposal is that a reason causes belief by indicating to the subject that some propositional content is true. There are seemingly only two ways in which this can occur. Either one is apparently aware of a support relation between some other proposition, which is taken to be true, and this new content. Or else one is apparently aware of the truth of the proposition itself. In both cases, the most proximate causes of the belief are phenomenal states. If our beliefs were not caused by phenomenal states, then they would not be *rationally formed* beliefs at all. They would result not from rational belief forming dispositions but something more like belief reflexes. Furthermore, the phenomenal properties of such states are sufficient to immediately cause belief in us. Any other properties of these states are superfluous to explaining why belief was rationally formed. Thus, there is a strong case that the proximate causes of our rationally formed beliefs within a fully-articulated, entirely non-superfluous folk-psychological explanations are always phenomenally individuated states.

What all this suggests is that, on causal internalism, justification ends up supervening on our phenomenally individuated mental states, or some subset of them. If this is correct, then causal internalism implies (materially if not strictly) some species of what Declan Smithies calls “phenomenalism”.[[52]](#footnote-52) What makes phenomenally individuated states so important, however, is not that they are reflectively transparent to us (as Smithies would have it), but that they are the only ones directly accessible to our rational belief forming systems.[[53]](#footnote-53) It is not their transparency that distinguishes them, but their causal role. More precisely, it is the fact that phenomenally individuated states can play the unique causal role assigned to *reasons*: causing belief, yes, but doing so by indicating the truth of its content. Plausibly, such indication is only possible because of our first-order awareness (or potential awareness) of those states.

 The upshot is that, on causal internalism, it is plausible that all of the contributors to justification are ones we are either aware of or potentially aware of on the first-order level. If so, then causal internalism manages once again to capture a central element of access internalism without taking on its excesses.

**Works Cited**

Bergmann, Michael. *Justification Without Awareness: A Defense of Epistemic Externalism*. Oxford, UK: Oxford University Press, 2006.

Berker, Selim. “Luminosity Regained.” *Philosophers’ Imprint* 8 (2008): 1–22.

BonJour, Laurence. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.

BonJour, Laurence, and Ernest Sosa. *Epistemic Justification: Internalism Vs. Externalism, Foundations Vs. Virtues*. Hoboken, NJ: Wiley-Blackwell, 2003.

Carruthers, Peter. *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford, UK: Oxford University Press, 2011.

Chisholm, Roderick M. *Theory of Knowledge*. 2nd edition. Englewood Cliffs, N.J.: Prentice-Hall, 1977.

Conee, Earl, and Richard Feldman. *Evidentialism: Essays in Epistemology*. Vol. 48. Oxford, UK: Oxford University Press, 2004.

Fumerton, Richard. *Epistemology*. Wiley-Blackwell, 2006.

Ginet, C. *Knowledge, Perception and Memory*. Philosophical Studies Series. Heidelberg, DE: Springer Netherlands, 1975.

Goldman, Alvin I. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press, 1986.

———. “What Is Justified Belief?” In *Justification and Knowledge*, edited by George Pappas, 1–23. Dordrecht, NL: D. Reidel, 1979.

Gopnik, Alison. “How We Know Our Minds: The Illusion of First-Person Knowledge of Intentionality.” *Behavioral and Brain Sciences* 16, no. 1 (1993): 1–14.

Huemer, Michael. “Compassionate Phenomenal Conservatism.” *Philosophy and Phenomenological Research* 74, no. 1 (2007): 30–55.

Langsam, Harold. “Mcdowell’s Infallibilism and the Nature of Knowledge.” *Synthese* 198, no. 10 (2021): 9787–9801.

McAllister, Blake. “Justification Without Excuses: A Defense of Classical Deontologism.” *American Philosophical Quarterly*, forthcoming.

Owens, David. *Reason Without Freedom: The Problem of Epistemic Normativity*. Abington, UK: Routledge, 2000.

Plantinga, Alvin. *Warrant: The Current Debate*. Oxford, UK: Oxford University Press, 1993.

Pollock, John. *Contemporary Theories of Knowledge*. 1st Edition. London, UK: Hutchinson, 1986.

———. “Epistemic Norms.” *Synthese* 71, no. 1 (1987): 61–95.

Pollock, John, and Joseph Cruz. *Contemporary Theories of Knowledge*. Second Edition. Washington DC, USA: Rowman & Littlefield, 1999.

Schoenfield, Miriam. “Internalism Without Luminosity.” *Philosophical Issues* 25, no. 1 (2015): 252–72.

Schwitzgebel, Eric. *Perplexities of Consciousness*. Denver, CO: Bradford, 2011.

———. “The Unreliability of Naive Introspection.” *Philosophical Review* 117, no. 2 (2006): 245–73.

Smithies, Declan. *The Epistemic Role of Consciousness*. Oxford, UK: Oxford University Press, 2019.

———. “The Phenomenal Basis of Epistemic Justification.” In *New Waves in Philosophy of Mind*, edited by Jesper Kallestrup and Mark Sprevak, 98–124. London, UK: Palgrave MacMillan, 2014.

Srinivasan, Amia. “Normativity Without Cartesian Privilege.” *Philosophical Issues* 25, no. 1 (2015): 273–99.

Steup, Matthias. “The Deontic Conception of Epistemic Justification.” *Philosophical Studies* 53, no. 1 (1988): 65–84.

Wedgwood, Ralph. “Internalism Explained.” *Philosophy and Phenomenological Research* 65, no. 2 (2002): 349–69.

Williamson, Timothy. *Knowledge and Its Limits*. Oxford, UK: Oxford University Press, 2000.

1. Plantinga 1993, 15. [↑](#footnote-ref-1)
2. See McAllister forthcoming for a defense of this account of epistemic blame and its connection to justification as originally conceived. [↑](#footnote-ref-2)
3. See, e.g., Littlejohn forthcoming, Srinivasan 2015, Williamson forthcoming. [↑](#footnote-ref-3)
4. This is the account of Descartes and Locke (Plantinga 1993, Ch. 1). Some contemporary proponents of this account (or key elements of it) include Bonjour 1985, Chisholm 1977, Ginet 1975, Steup 1988, and Smithies 2019. See also Srinivasan 2015, §3, for discussion. [↑](#footnote-ref-4)
5. On the definition of reflective transparency, see Williamson 2000 and Berker 2008. [↑](#footnote-ref-5)
6. Huemer 2007, 35; Srinivasan 2015, 275. [↑](#footnote-ref-6)
7. Gopnik 1993; Carruthers 2011; Schwitzgebel 2006, 2011. [↑](#footnote-ref-7)
8. Williamson 2000, Ch. 4. [↑](#footnote-ref-8)
9. Kvanvig 2014, Ch. 2, §5. [↑](#footnote-ref-9)
10. Smithies 2019. [↑](#footnote-ref-10)
11. Srinivasan 2015, 273. [↑](#footnote-ref-11)
12. See Schoenfield 2015 for a related attempt to motivate internalism despite threats to luminosity. There are significant differences in how we go about formulating and defending our ideas, nevertheless there is, I think, a common insight motivating both of our responses. [↑](#footnote-ref-12)
13. Pollock 1987, 64; 1986, Ch. 5. [↑](#footnote-ref-13)
14. Pollock 1987, 66. [↑](#footnote-ref-14)
15. If you think that perceptual beliefs are inferred from how things appear to us, then consider how it is we came to believe that things appear thus-and-so. At some point, our beliefs must be caused apart from higher-order beliefs, or else we would be faced with forming an infinite regress of beliefs prior to forming any belief and belief formation would be impossible. [↑](#footnote-ref-15)
16. Williamson 2000, 180; see also Owens 2000, 22. [↑](#footnote-ref-16)
17. This does not yet commit us to the position that those appearances are *good* reasons for belief, or ones adequate to justify belief in the absence of defeaters. [↑](#footnote-ref-17)
18. Pollock 1987, 68. [↑](#footnote-ref-18)
19. With respect to the norms of justification, a belief that conforms to those norms is said to have propositional justification for its content, whereas a belief that complies with those norms is said to be doxastically justified. [↑](#footnote-ref-19)
20. Pollock 1987, 69. [↑](#footnote-ref-20)
21. Wedgwood 2002, 357. [↑](#footnote-ref-21)
22. *Ibid.* [↑](#footnote-ref-22)
23. The fact that the man heard the dryer go off on a Wednesday couldmake a difference to what he believes if, for instance, he knows that the dryer goes off sporadically (even when the laundry isn’t finished) on days other than Wednesday. However, the man would have to be reflectively aware of the fact that he heard the dryer on a Wednesday in order for it factor into belief-formation, in which case its being Wednesday doesn’t make any difference to whether the sound of the dryer immediately brings about belief or not. Another way to put this is that the state *hearing the dryer go off on a Wednesday* is, in that particular case, not the most proximate cause of belief, but only causes belief by virtue of intermediary states such as believing that it is Wednesday. [↑](#footnote-ref-23)
24. Wedgwood 2002, 363. [↑](#footnote-ref-24)
25. Conee and Feldman 2004, Ch. 3. [↑](#footnote-ref-25)
26. Bergmann 2006, 55-56. [↑](#footnote-ref-26)
27. If we define “externalism” as “not-causal internalism”, then it follows by definition that externalist norms appeal to conditions that are not directly accessible to our automatic processing systems. That being said, I should showcase that the sorts of theories usually thought of as exemplars of externalism are still “externalist” in this particular sense. [↑](#footnote-ref-27)
28. See, e.g., Goldman 1979. [↑](#footnote-ref-28)
29. It is of course possible to gerrymander an external state of affairs that necessarily correlates with some internal state (e.g., the state of affairs *S has internal state M and God exists* is an external state of affairs that, if God necessarily exists, will perfectly track the internal state M), but these are trivial. The crucial thing is that none of the external conditions proposed by externalist theories of justification, nor anything in their vicinity, are of this sort. [↑](#footnote-ref-29)
30. Pollock 1987, 74. [↑](#footnote-ref-30)
31. See, e.g., Goldman 1986. [↑](#footnote-ref-31)
32. Wedgwood 2002, §4. [↑](#footnote-ref-32)
33. One of these might feature in a folk-psychological explanation of belief if, say, S1 believes that trusting in the testimony of S2 is a reliable process, and so infers the belief that p from the fact that S2 testified that p. But clearly the reliability of that process is not the *most proximate* cause of belief in that situation. [↑](#footnote-ref-33)
34. If the content of their beliefs is externally determined, then perhaps the subjects do not form the exact same beliefs. Nevertheless, it is clear that one subject will form the same kind of belief if and only if the other does. The only difference is that the content of the one belief will be “filled in” differently than the content of the other. Thus, I will set this complication aside. [↑](#footnote-ref-34)
35. Huemer 2007, 40. [↑](#footnote-ref-35)
36. Schoenfield 2015, 257-258. I’ve amended the example to better fit our purposes. [↑](#footnote-ref-36)
37. Williamson 2000, 180. [↑](#footnote-ref-37)
38. Langsam 2021, 9792. [↑](#footnote-ref-38)
39. Williamson 2000, 180. [↑](#footnote-ref-39)
40. Langsam 2021, 9799. [↑](#footnote-ref-40)
41. Futhermore, nothing said above about the causes of our beliefs implies that we can have the same token state in a case of perception as we do in a case of illusion or hallucination. Say that, in the case of perception, S has a perceptual state P and, in the case of hallucination, S has an entirely different token state H. There is no token state common to both of them. The argument made above is simply that P would be the most proximate cause of a belief in S if and only if H would be as well (were S to have that state instead), and this proves that P only brings about belief by virtue of properties that H also possesses. Now, if we wish, we could define “appearances” as a type of mental state characterized exclusively by how it feels to be in a state like P or H, in which case we could say that S has an appearance in both the perceptual case and the hallucinatory case; and those appearances would be the proximate causes of S’s belief in both. But it would not be *the same* appearance in both cases, only the same *type* of appearance. And it is only the former claim that disjunctivists should be concerned to deny. Thus, once again, we see that even disjunctivists should accept the causal claims at issue here. [↑](#footnote-ref-41)
42. Schoenfield 2015, 254. [↑](#footnote-ref-42)
43. Being the proximate cause of a belief is not sufficient for being the base of that belief, but the additional conditions (whatever they are) are ones that will be ordinarily satisfied when a directly accessible mental state triggers our rational belief forming dispositions. [↑](#footnote-ref-43)
44. Bergmann 2007, 70. [↑](#footnote-ref-44)
45. Something like this is perhaps by Bergmann does not see internalism and externalism as exhaustive, with mentalism (including what we are calling “causal internalism”) residing somewhere in between the two. [↑](#footnote-ref-45)
46. The claim under consideration here adds to Wedgwood 2002 that the internal facts about our mental states that determine belief formation always have to do with their phenomenal properties. [↑](#footnote-ref-46)
47. Smithies 2019, 25. See also Smithies 2014. [↑](#footnote-ref-47)
48. Bonjour and Sosa 2003, 4.2-4.3. [↑](#footnote-ref-48)
49. Bergmann 2006, 67. [↑](#footnote-ref-49)
50. Wedgwood 2002, 359-360. [↑](#footnote-ref-50)
51. Once again, issues of content externalism aside. [↑](#footnote-ref-51)
52. Smithies 2019. [↑](#footnote-ref-52)
53. This cuts against Smithies claim that phenomenalism and a phenomenal conception of evidence is best motivated by *accessibilism*—the position that justification is reflectively transparent to us. This claim features centrally in the argument of Smithies 2019, and especially in his critique of phenomenal conservatism in Ch. 12. Essentially, Smithies proposes that internalists double-down on reflective transparency, whereas I am trying to show how internalism can be motivated apart from reflective transparency. [↑](#footnote-ref-53)