DEFEATERS AND HIGHER-LEVEL REQUIREMENTS

BY MICHAEL BERGMANN

Internalists tend to impose on justification higher-level requirements, according to which a belief is justified only if the subject has a higher-level belief (i.e., a belief about the epistemic credentials of a belief). I offer an error theory that explains the appeal of this requirement: analytically, a belief is not justified if we have a defeater for it, but contingently, it is often the case that to avoid having defeaters, our beliefs must satisfy a higher-level requirement. I respond to the objection that externalists who endorse this error theory will be forced to accept a radical form of scepticism.

In this paper I shall be developing a natural view of defeaters according to which it is analytic that a belief is not justified if it has a defeater. Then I shall use this view to account for the attractiveness of the common but mistaken thesis that there is a higher-level requirement on justification.

Internalist theories of epistemic justification tend to require for justification of a belief b that anyone holding b must have a certain perspective on it, must conceive of b as having something or other going for it. For example, the internalist might require for b’s justification that the believer must believe that b is formed in a reliable way. This is called a ‘higher-level requirement’ because it requires, for justification, a higher-level belief, i.e., one about a belief’s epistemic credentials. If one understands this higher-level requirement as perfectly general, so that the required higher-level belief must itself be justified, then the requirement gives rise to a regress viewed by many as vicious.1 Hence a popular objection to some versions of internalism is that they are committed to higher-level requirements which lead to vicious regresses.

Supposing the regress is vicious, what should we conclude about this higher-level requirement? One option is simply to deny that there is any such requirement on justification. The problem with this move is that it does

not do justice to the *prima facie* plausibility of the suggestion. Another option, attractive to those convinced that the requirement is a plausible one, is to say that it is required for the justification of object-level beliefs (i.e., beliefs that are not about a belief’s epistemic credentials), but not for the justification of all higher-level beliefs. The problem with this response to the regress problem is that it seems to be *ad hoc*. What plausible reason is there for thinking that there is a higher-level requirement for the justification of object-level beliefs (and perhaps some higher-level beliefs) which is not an equally plausible reason for thinking that there is a higher-level requirement for the justification of all beliefs?

In §II of this paper, I propose a third option which, like the first, says that there is no higher-level requirement on justification for beliefs at any level, but like the second option, acknowledges that there is something to the idea that there is such a requirement. What this has going for it is that it is often *contingently* the case that in order for beliefs to avoid having defeaters, they must satisfy a higher-level requirement. In §III, I consider and respond to an objection which arises for externalists who want to adopt my §II account of why the higher-level requirement on justification seems mistakenly so appealing to internalists. According to that objection, my account has sceptical implications which externalists will want to reject. But first, in §I, I lay out some key notions to be employed in the paper, the most important of which is the notion of a defeater.

I. DOXASTIC PROPOSITIONAL ATTITUDES AND DEFEATERS

In giving my account of why one might think that there is a higher-level requirement on justification I consider various doxastic propositional attitudes which one might take towards certain higher-level propositions. And I ask whether any of those attitudes constitute defeaters for the beliefs those higher-level propositions are about. I shall begin, therefore, by saying something about how I shall understand doxastic propositional attitudes and defeaters.

(a) Doxastic propositional attitudes

Three attitudes one might take towards a proposition \( p \) are believing \( p \), disbelieving \( p \) (i.e., believing \( p \) is false), and withholding \( p \) (i.e., refraining from either believing or disbelieving \( p \)). Is it possible to take none of these three

---


attitudes towards \( p \)? Certainly. One can take no attitude whatsoever towards \( p \) (not even withholding it). One way in which this may happen is if one never even considers \( p \). Indeed, it is natural to think that we take no attitude towards most propositions, because we have never considered them.

(Although I do not think that taking an attitude towards a proposition \( p \) requires considering \( p \), I do think it requires having once considered it. Thus I follow Alvin Goldman in making a distinction between being disposed to believe \( p \) and having a dispositional belief that \( p \). To have an occurring belief that \( p \) involves considering \( p \). To have a dispositional belief that \( p \) involves, at the very least, having stored in memory an affirming attitude towards \( p \), which you once considered and at that time originally affirmed. But one can have a disposition to believe \( p \) without believing \( p \), either occurringly or dispositionally. For example, many people are, no doubt, disposed to believe that zebras do not wear sports coats in the wild, even though they do not in fact believe that proposition, either occurringly or dispositionally, since they have never considered it.)

Withholding \( p \), then, is a propositional attitude distinct from mere failure to take up any attitude towards \( p \). Like believing or disbelieving, it is taking an attitude towards a proposition. What more can one say about withholding? As I shall be using the term, withholding \( p \) involves resistance, voluntary or involuntary, to believing \( p \) and to disbelieving \( p \). The only thing one must consider in order to believe \( p \) or to disbelieve \( p \) is \( p \) (or its denial). But to withhold \( p \) (in the sense I have in mind) one must, in addition, consider the prospect of one’s believing \( p \) as well as the prospect of one’s disbelieving \( p \); otherwise one will not be able to resist both believing \( p \) and disbelieving \( p \). So withholding \( p \) involves not only an attitude towards \( p \) but also attitudes towards attitudes towards \( p \). I shall call these three attitudes, i.e., believing, disbelieving and withholding, as I have described it here, ‘the doxastic attitudes’. Together with the possibility of taking no doxastic attitude at all, this gives four ways of relating to a proposition.

Is it possible to take none of the three doxastic attitudes towards \( p \) while at the same time taking some propositional attitude or other towards \( p \) (such as hoping or fearing or imagining or considering that \( p \))? Here again I think the answer is ‘Yes’. Considering that \( p \) is compatible with each of the three doxastic attitudes towards \( p \). But it is possible to consider \( p \) without believing or disbelieving \( p \) (the possibility of withholding guarantees this). And it is possible for someone who neither believes nor disbelieves \( p \) to consider \( p \) without giving a thought to whether believing or disbelieving it is appropriate (and so with no attitude of resistance towards the prospect of believing or withholding \( p \).)

---

One might consider $p$ with the purpose of determining whether it really is a proposition, and be so distracted by that purpose as to lack any of the doxastic attitudes to it. So there are two ways of taking no doxastic attitude towards $p$: by taking no attitude at all towards $p$ or by taking only a non-doxastic attitude towards $p$.

(b) Defeaters

For any proposition $p$ and any person $S$, I shall call ‘$p^*$’ the proposition that $S$’s belief-that-$p$ is formed in a reliable way. A question I shall address in explaining the appeal of higher-level requirements is this: are any of the four ways in which $S$ can be related to $p^*$ defeaters for $S$’s belief that $p$?

To answer this question I need to distinguish between what I shall call ‘propositional’ defeaters (which are propositions) and ‘mental state’ defeaters (which are either propositional attitudes or experiences or combinations thereof). According to defeasibility accounts of knowledge, a propositional defeater for a belief $b$ is a true proposition (typically, one not believed by the person holding $b$), the mere truth of which prevents $b$ from counting as knowledge. Thus defeasibility accounts add to ‘justified true belief’ accounts the requirement that there must be no true proposition which is a propositional defeater for the belief in question. What this amounts to, roughly, is that there is no true proposition such that if it were added to the subject’s evidence base, the belief in question would not be justified. But I shall rather focus on mental state defeaters (henceforth, ‘defeaters’), which for my purposes can be defined as follows:

$D_1.$ $d$ is a defeater at $t$ for $S$’s belief $b$ iff (i) $d$ is an experience or propositional attitude or combination thereof; (ii) $S$ comes to have $d$ at $t$; (iii) as a result of $S$’s coming to have $d$ at $t$, $b$ ceases to be justified.

Given this account of a defeater, it becomes analytic (and therefore completely uncontroversial) that a belief is justified only if there is no defeater for it. Thus according to ($D_1$), if the defeated power of a defeater $d$ is neutralized or defeated, $d$ ceases to be a defeater. ($D_1$) is an account of what one might call a ‘full defeater’. For an account of a partial defeater, clause (iii) of ($D_1$) must be changed so as to say that as a result of $S$’s coming to have $d$, $b$ becomes less justified.

This account (like Plantinga’s) is really only an account of one kind of newly acquired defeater, acquired by coming to have the propositional attitude or experience which is the defeater. There are also the kind of newly acquired defeaters where a propositional attitude or experience one

already has comes to have newly acquired defeating power. And in addition there are defeaters that are not newly acquired. However, because it is easier to explain what newly acquired defeaters consisting of newly acquired propositional attitudes or experiences are, and because this will suffice for my purposes in this paper, I shall not attempt here to give a more general account of defeaters.

Why say, as (D1) does, that \( d \) is a defeater for \( b \) only if the result of coming to have \( d \) is that \( b \) ceases to be justified? Is it not enough (for \( d \) to result in \( b \)'s having a defeater) that the subject takes \( b \) to lose its justification as a result of coming to have \( d \)? Yes, it is enough. But (D1) already handles this. If \( x \) is taking the result of one’s coming to have \( d \) to be that \( b \) ceases to be justified, and \( 'd^* \) stands for the combination of \( x \) and \( d \), then the question is whether \( d^* \) gives one a defeater for \( b \). The answer is that it does. But the reason why it does is that \( b \) in fact ceases to be justified as a result of \( S \)'s coming to have \( d^* \). This can be put in terms of a distinction between two kinds of mental state defeaters for \( b \): believed defeaters, which are propositional attitudes or experiences \( S \) has which \( S \) believes make \( b \) unjustified, and actual defeaters, which are propositional attitudes or experiences \( S \) has which in fact make \( b \) unjustified. My claim, in response to the proposal that actual defeaters are not necessary for defeating justification (because believed defeaters seem to be sufficient for that), is that all believed defeaters \( S \) has for \( b \) are actual defeaters \( S \) has for \( b \), though not vice versa. (Below, in §1(c), I respond to some objections to the claim that all believed defeaters are actual defeaters.)

(It might be helpful to mention here how this distinction between believed and actual defeaters applies to a necessary condition of warrant, which I have called the ‘no-defeater condition’ (NDC), and which is featured prominently in some of my other work.\(^6\) The sort of defeater that according to (NDC) must be absent is a believed defeater. This clarification is important, because although it is analytic that a no-actual-defeater condition is necessary for warrant, it is a substantive position that a no-believed-defeater condition is necessary for warrant. Thus the view that (NDC), as I have been understanding it in the other writings mentioned above, is necessary for warrant is a somewhat controversial, though widely endorsed, position, not analytic. See my ‘Internalism, Externalism, and the No-Defeater Condition’, pp. 405–7, for evidence that both internalists and externalists seem to endorse the necessity for warrant of a no-believed-defeater condition.)

‘Rebutting’ and ‘undercutting’ defeaters can be distinguished as follows:

D2. \(d\) is a rebutting defeater for \(b\) iff \(d\) is a defeater for \(b\) which is (or is an epistemically appropriate basis for) the belief that \(b\) is false

D3. \(d\) is an undercutting defeater for \(b\) iff \(d\) is a defeater for \(b\) which is (or is an epistemically appropriate basis for) the belief that one’s actual ground or reason for \(b\) is not indicative of \(b\)’s truth.

For example, the belief that if God existed, he would not permit the sorts of evil that I in fact see around me is a rebutting defeater for my belief that God exists; an undercutting defeater for this belief would be a good reason for thinking that the actual ground for it is an unreliable Freudian sort of wish-fulfilment. The rebutting defeater gives me a reason to be an atheist, whereas the undercutting defeater gives me a reason to be an agnostic.

(c) Which doxastic attitudes are defeaters?

I asked above whether any of the four ways mentioned in §I(a) in which \(S\) can relate to \(p\) (the proposition that \(S\)’s belief that \(p\) is formed in a reliable way) are defeaters for \(S\)’s belief that \(p\). First, is \(S\)’s disbelieving \(p\) (i.e., \(S\)’s believing that \(p\) is false) a defeater for \(S\)’s belief that \(p\)? It seems so. If I come to believe that my belief that \(p\) is not formed in a reliable way, then my belief that \(p\) ceases to be justified.

But what if my belief that \(p\) is not formed in a reliable way is itself unjustified? Can a justified belief of mine lose its justification as a result of my coming to hold an unjustified belief? I think it can. It is clear that it is an epistemically bad state of affairs if I believe both that \(p\) and that my belief that \(p\) is not formed in a trustworthy way. Thus if one has a justified belief that \(p\) and then comes to have an unjustified belief that one’s belief that \(p\) is not formed in a reliable way, one will be in an epistemically bad state of affairs. It is not merely that it is epistemically bad to have the unjustified belief. There is the additional problem of believing that a belief of yours was formed in an inappropriate way.

Here are two questions:

Q1. What happens to the justification of \(S\)’s justified belief that \(p\) when \(S\) comes to have an unjustified belief that his belief that \(p\) is formed in an unreliable way?

Q2. When \(S\) has a justified belief that \(p\) and then comes to have an unjustified belief that his belief that \(p\) is formed in an unreliable way, what is the most epistemically appropriate way for the situation to change so that \(S\) can escape this epistemically bad state of affairs?
It seems clear that the answer to (Q2) is that the best way to escape this epistemically bad state of affairs is to give up the unjustified belief that the belief that \( p \) is formed in an unreliable way. But this answer to (Q2) does not suggest that the belief that \( p \), when held in conjunction with the belief that one’s belief that \( p \) is formed in an unreliable way, remains justified. For it seems perfectly natural and sensible to say that if you think your belief that \( p \) is formed in an unreliable way, it is unreasonable for you to believe that \( p \). It seems unreasonable to hold a belief you think was formed in an unreliable way, just as it seems unreasonable to hold a belief you think is false. It is true that if you have lots of evidence for \( p \) and none for the belief that \( p \) is unreliably formed, then part of your total relevant evidence strongly supports \( p \). But a belief’s justification is not determined by considering whether a part of the believer’s evidence supports it. One must look at the total relevant evidence. And for the case to be of the sort I have been describing, the total relevant evidence will include not only the ample evidence supporting \( p \) but also your further belief that this evidence, on which your belief that \( p \) was based, is not a reliable indicator of the truth of \( p \). If you think the evidence on which your belief that \( p \) was based is unreliable, then your belief that \( p \) is not justified, even if the evidence is in fact excellent evidence for \( p \). Good reasons or evidence are useless to you as justification, if you think (even unjustifiably or mistakenly) that they constitute terrible evidence.

But can this be correct? Does merely thinking one has a defeater give one a defeater? Is it true, as I said earlier, that all believed defeaters are actual defeaters? Does that not make it too easy to lose justification? Do philosophy students who are taken in by a professor’s convincing presentation of a sceptical argument really lose their justification for their perceptual beliefs? I think it depends. If the students seriously think their perceptual faculties are not to be trusted, then it seems that their perceptual beliefs are not justified. If on the other hand they remain convinced that their perceptual faculties are to be trusted, but are puzzled by the argument and so claim that they think their perceptual beliefs are not justified, their perceptual beliefs remain justified. What matters is whether one really thinks one’s faculties and evidence are to be trusted, not what one merely says about them.

We can make all these claims even though we think that the best thing to do, in the example two paragraphs back, is to stop thinking that one’s belief that \( p \) is formed in an unreliable way (which would make it possible for one’s belief that \( p \) to be justified). The temptation to conclude that an unjustified belief cannot defeat the justification of a justified belief arises from focusing on (Q2) above and on the answer I gave to it. But consideration of (Q1) suggests that S’s belief that \( p \) loses its justification if S gets into the epistemically bad state of affairs of believing (justifiedly or not) that the belief
was formed in an unreliable way. That loss of justification for the belief that \( p \) is at least one component of what the ‘badness’ of this epistemically bad state of affairs consists in.

Why think that unjustified defeating beliefs remove justification, if unjustified supporting beliefs cannot confer justification? Is that not arbitrarily to treat reasons against a belief more leniently than reasons for it? No. It would be ludicrous to say that unjustified beliefs could confer justification onto other beliefs they support. That would amount to saying that justification for \( b_2 \) can result solely from an inference from \( b_1 \) even if \( b_1 \) has no justification it can transmit. But nothing like that is being suggested when one says that unjustified defeating beliefs can remove justification: here there is no claim that an inference can transfer justification that is not there to be transferred. Instead, the point is just that it is epistemically bad for one’s belief \( b \) if one comes to have a belief (justified or not) that \( b \) is formed in an unreliable way. And, for the reasons noted above, it is plausible to think that this epistemic badness can cause \( b \) to lose its justification. It therefore seems perfectly reasonable, consistent and non-arbitrary to think that unjustified beliefs can defeat justification even though they cannot confer it.

So disbelieving \( p^* \) is a defeater for \( S \)’s belief that \( p \), even when this disbelief is unjustified. What about withholding \( p^* \)? Is that a defeater for \( S \)’s belief that \( p \)? Suppose Sally is in a factory’s viewing room, looking through a safety window at a machine manufacturing widgets, and forms the belief (based on how they look to her) that the widgets on the conveyor belt are red. Someone asks her ‘Are those widgets red, or do they just look red because there is a red light shining on them?’. Sally, like the questioner, has no idea whether there is a red light shining on the widgets; she does not even know how likely it is that there would be. Being a reflective person, Sally now considers the higher-level proposition that her belief \( The\ widgets\ are\ red \) is formed in a reliable way. Being completely uncertain about whether that higher-level proposition is true, she resists believing both it and its denial. In other words, if \( p \) is the proposition \( The \ widgets \ are \ red \), she withholds \( p^*_{Sally} \). Does this give her, in these circumstances, a defeater for her belief that the widgets are red? I think it does. If you are considering whether the actual basis of your belief that \( p \) is indicative of \( p \)’s truth and you find yourself resisting the belief that it is (because you have considered the matter and you have no idea whether it supports \( p \) or not), that seems to undercut your justification for believing \( p \) in the same way as if you believed outright that the actual basis for your belief that \( p \) did not indicate \( p \)’s truth. And for reasons similar to those discussed above in connection with believing that your belief that \( p \) is unreliably formed, this result holds even if your attitude of uncertainty about whether your belief that \( p \) is formed in a reliable way
is unjustified. What this suggests is that definition (D3) should be modified as follows:

\[ \text{D3*} \text{. } d \text{ is an undercutting defeater for } b \text{ iff } d \text{ is a defeater for } b \text{ which (i) is (or is an epistemically appropriate basis for) the belief that one's actual ground or reason for } b \text{ is not indicative of } b \text{'s truth, or (ii) is (or is an epistemically appropriate basis for) an attitude of significant uncertainty about the proposition that one's actual ground or reason for } b \text{ is indicative of } b \text{'s truth, significant enough to withhold that proposition.} \]

Given this definition, what I have said above suggests that S’s withholding \( p^*s \) is an undercutting defeater for S’s belief that \( p \). Later I shall consider reasons to question this, without questioning (D3*). But for now I need to have before us the sort of reasoning that makes the suggestion plausible.

II. THE APPEAL OF HIGHER-LEVEL REQUIREMENTS

(a) Why higher-level requirements are appealing

Suppose that as suggested at the end of the previous section, S’s disbelieving \( p^*s \) and S’s withholding \( p^*s \) are both defeaters for S’s belief that \( p \). This will suggest, to those who think that one’s only options with respect to any proposition are to believe it, disbelieve it or withhold it, that to avoid having a defeater for his belief that \( p \), S must believe \( p^*s \). This thought is a very natural one, simply because it is natural not to consider the option of taking none of the three doxastic attitudes towards \( p^*s \); and this provides at least a partial explanation of why the view that there is a higher-level requirement on justification is so appealing.

But the idea that there is a higher-level requirement on justification may still seem plausible even if one recognizes the option of taking no doxastic attitude towards a proposition. If S’s disbelieving \( p^*s \) and withholding \( p^*s \) are defeaters for S’s belief that \( p \), then in order for S’s belief that \( p \) to avoid being defeated, S must either hold the higher-level belief that \( p^*s \) or take no doxastic attitude whatsoever towards \( p^*s \). Thus given that a belief is justified only if it is not defeated, we may conclude that

1. For any person S and any proposition p which S believes, if S takes up some doxastic attitude towards \( p^*s \), then S’s belief that \( p \) is justified only if S has the higher-level belief that \( p^*s \).

Now consider

2. S takes up some doxastic attitude towards \( p^*s \).
If one thought that for any person $S$ and any proposition $p$ that $S$ believes, (2) is true, then one could conclude that for any person $S$ and any proposition $p$,

3. $S$'s belief that $p$ is justified only if $S$ has the higher-level belief that $p^*$.  

Or, if one thought that (2) was true for most subjects and all (or most) propositions they believe, then one could conclude that (3) is true for most subjects and all (or most) propositions they believe.

But of course the suggestion that (2) is true for all (or most) subjects, and all (or most) propositions they believe, is not very plausible (especially not if the propositions in question are themselves higher-level propositions). Furthermore, even if this suggestion were plausible, this would (if true) be merely a contingent matter of fact. What is required for justification is the absence of a defeater. And if it just so happened that as a contingent matter of fact, (2) was true for all subjects and all propositions they believe, it would follow that people can avoid having a defeater for their beliefs only if those beliefs satisfy a higher-level requirement. This might explain why some philosophers are tempted to think there is a higher-level requirement on justification; but it does not show that there is such a requirement. It establishes at most that as a contingent matter of fact, our beliefs avoid being defeated only if (and therefore are justified only if) they satisfy a higher-level requirement. It is the absence of a defeater, not the satisfaction of a higher-level requirement, that is essential for justification.

But suppose that a defender of higher-level requirements said not that (2) is in fact true for most or all subjects and most or all propositions they believe, but rather that (2) is a necessary condition of the justification of $S$'s belief that $p$. From this together with (1) it would follow that there is a higher-level requirement on justification (assuming (1) is a necessary truth). Even if it is not explicitly rehearsed, it is the attractiveness of exactly this sort of reasoning that explains (at least partially) the appeal of higher-level requirements on justification.

One noteworthy example of an internalist who seems to rely, even if only implicitly, on this sort of reasoning in defending a higher-level requirement on justification is BonJour (pp. 41–3), in his discussion of his well known objection to externalism involving the clairvoyant Norman. BonJour’s goal is to persuade us to conclude, about a case in which all externalist conditions on justification are satisfied, that the subject’s belief is unjustified. He first considers examples of people with reliable clairvoyant beliefs who have defeaters for their beliefs (such as evidence for the conclusion that there is no such thing as clairvoyance, or evidence for the falsity of the particular clairvoyant belief formed). But he thinks an externalist will simply agree that because these beliefs have defeaters, these are not cases of justified belief.
And since an externalist can consistently require for justification the absence of a defeater, this prevents these cases from counting as objections to externalism. So BonJour (p. 41) turns to the case of Norman:

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

BonJour thinks it is clear that Norman lacks justification, despite satisfying both a reliabilist condition and a no-actual-defeater condition.

I shall not consider here the merits of this example as an objection to externalism. Instead, I shall focus on what BonJour says in discussing it. He notes (p. 42) that in the above description of the case, he neglects to mention whether Norman ‘believes himself to have clairvoyant power’. And so he asks us to consider both alternatives: that Norman does believe this, and that he does not. When considering the first of these alternatives, BonJour says (ibid.) that Norman would clearly be unjustified in believing he has clairvoyant power, because he has no evidence or even any externalist justification for belief in that proposition (though BonJour does not explain why Norman lacks externalist justification for belief in that proposition). Here is what he says (p. 42) when considering the second alternative (i.e., that Norman does not believe himself to have clairvoyant power):

... this would mean that Norman must not believe that he has the power of clairvoyance (or at least that his acceptance of that belief about the President’s whereabouts must not depend on his having such a belief). But if this specification is added to the case, it becomes quite difficult to understand what Norman himself thinks is going on. From his standpoint, there is apparently no way in which he could know the President’s whereabouts. Why then does he continue to maintain the belief that the President is in New York City? Why isn’t the mere fact that there is no way, as far as he knows, for him to have obtained this information a sufficient reason for classifying this belief as an unfounded hunch and ceasing to accept it? And if Norman does not do this, isn’t he thereby being epistemically irrational and irresponsible?

Thus, I submit, Norman’s acceptance of the belief about the President’s whereabouts is epistemically irrational and irresponsible, and thereby unjustified, whether or not he believes himself to have clairvoyant power, so long as he has no justification for such a belief. Part of one’s epistemic duty is to reflect critically upon one’s beliefs,

3 See my ‘Internalism, Externalism, and the No-Defeater Condition’, pp. 405–7, for evidence that both internalists and externalists acknowledge that imposing such a requirement is consistent with being an externalist, and ‘A Dilemma for Internalism’, §4, for an explanation of why they are consistent.
and such critical reflection precludes believing things to which one has, to one’s knowledge, no reliable means of epistemic access.

If \( p \) is the proposition \textit{The President is in New York City} and \( N \) is Norman, then clearly BonJour’s claim is that if Norman does not believe \( p^* \), then his belief that \( p \) is not justified.

What reasons does BonJour give for this claim? It seems that he thinks that Norman must take up some doxastic attitude towards \( p^* \) in order for his belief that \( p \) to be justified. This, I take it, is the upshot of his talk, in the last sentence quoted, of ‘critical reflection’. And as I mentioned earlier, he thinks Norman would not be justified in taking the attitude of belief towards \( p^* \). This leaves only the doxastic attitudes of disbelief and withholding. But BonJour thinks, it seems, that taking either of those attitudes towards \( p^* \) results in a defeater for the belief that \( p \). This is why he thinks that satisfying both a reliability condition and a no-actual-defeater condition is insufficient for justification: it fails to take account of the consequences of the further requirement on justification according to which \( S \) must take up some doxastic attitude or other towards \( p^* \).

(b) \textit{Why the appeal of higher-level requirements is misleading}

My account of why higher-level requirements are appealing, then, is this. Internalists find it tempting to think that we ought to reflect critically on our beliefs, and hence that

\begin{itemize}
  \item \( S \) takes up some doxastic attitude towards \( p^* \)
\end{itemize}

is a necessary condition for the justification of \( S \)’s belief that \( p \). Likewise, they think that if \( S \) withholds or disbelieves \( p^* \), \( S \) has a defeater for the belief that \( p \). Thus they conclude that \( S \) must believe \( p^* \) in order to be justified in believing \( p \). Again I am not claiming that all who are influenced by this line of reasoning explicitly rehearse it before endorsing a higher-level requirement. Nevertheless, those who find higher-level requirements attractive seem to be assuming that if one’s belief does not satisfy a higher-level requirement, one must either question the reliability of one’s belief source or be irresponsible by not thinking at all about whether the source is reliable. (Some who endorse a higher-level requirement for justification might require not \( S \)’s belief that his belief that \( p \) is formed in a reliable way, but instead, say, his belief that it is adequately supported. For expository convenience, I have ignored possible alternative formulations of the higher-level requirement. But the argument above can easily be reformulated to allow for the alternatives.)

But is (2) a necessary condition for the justification of \( S \)’s belief that \( p \)? How could satisfying (2) possibly contribute to the justification? The idea is
that by considering and taking some doxastic attitude towards \( p^* \), S is thereby doing something that contributes to the justification of his belief that \( p \). But of course if the doxastic attitude S takes towards \( p^* \) is disbelief or withholding, this will not contribute to the justification of S’s belief that \( p \). So it is only taking the doxastic attitude of belief towards \( p^* \) that can possibly contribute to the justification of S’s belief that \( p \). But will just any belief, justified or not, do so? Would an irrational, irresponsible or insane belief that \( p^* \) contribute to the justification of S’s belief that \( p \)? No. Suppose S’s belief that \( p \) satisfies one’s own favourite external conditions and that S has no defeater for it and no doxastic attitude towards \( p^* \). Would the justification of S’s belief that \( p \) be increased if S had an irrational, irresponsible, insane or unjustified belief that \( p^* \)? Certainly not. So the only way in which (2) can sensibly be required for the justification of S’s belief that \( p \) is if we stipulate that it must be made true by S’s having a justified belief that \( p^* \).

But to require that (2) must be satisfied by S’s having a justified belief that \( p^* \) is to face the dilemma noted earlier in the paper: either this requirement is a general one applying to all beliefs, which leads to a vicious regress, or it applies to beliefs at the object-level and perhaps at some higher levels but not at all higher levels, an \textit{ad hoc} restriction made for no reason other than to avoid the regress. I conclude, therefore, that although it is contingently the case that some beliefs are justified only if the person holding them has some higher-level belief, because it is contingently the case that some beliefs can avoid having a defeater only by this means, there is no general or principled higher-level \textit{requirement} on justification.

### III. THE ‘HOUSE OF CARDS’ EFFECT AND SCEPTICISM

In §II I have given an error theory for those who find higher-level requirements on justification appealing. I have explained why such requirements only seem to apply to justification. This sort of error theory will be particularly attractive to externalists, since it is internalists who are most enamoured of such requirements and who make use of them in their objections to externalism.

But if one ponders the implications of the endorsement of this error theory by externalists, it might seem to result in a problem. For they may then be presented with the following objection: ‘Since you are an externalist, it is likely that you will reject the suggestion that most of our beliefs are unjustified. But by adopting the §II account of the appeal of higher-level
requirements on justification, you are forced to accept the sceptical conclusion that most or all of your beliefs are unjustified.' If this objection cannot be met, my error theory will not be attractive to either internalists (at least not those who are being charged with error) or externalists (on pain of being forced towards scepticism). In what follows, I shall first lay out the reasoning that underlies this sort of objection and then respond to it.

(a) The ‘house of cards’ effect

Suppose Sam, a philosophically reflective person who has read and accepted all I have said thus far, holds the object-level belief that \( p \), and thinks that his options for avoiding its defeat are (i) having the higher-level belief \( p^*_{Sam} \), or (ii) having no doxastic attitude at all towards \( p^*_{Sam} \). Because he thinks this, he finds himself unable to take option (ii). For he finds that while he is considering what I have said in this paper, he cannot help considering \( p^*_{Sam} \) and focusing on whether or not he believes it. The result is that he accepts \( p^*_{Sam} \). But being in a reflective mood, he finds himself again helplessly dwelling on the fact that he can avoid having a defeater for this only by (i) having the further higher-level belief that \( (p^*_{Sam})^*_{Sam} \), or (ii) having no doxastic attitude at all towards \( (p^*_{Sam})^*_{Sam} \). And given that he is dwelling on these matters, he finds once again that (ii) is not an option. So he forms the higher-level belief \( (p^*_{Sam})^*_{Sam} \). This reflective process can continue only so long before Sam fails to be able even to grasp the next higher-level proposition. However, suppose that before this occurs (in fact, at the very next level) he finds that, in considering \( ((p^*_{Sam})^*_{Sam})^*_{Sam} \), i.e., the proposition

My belief that \{my belief that \{my belief that \( p \) is formed in a reliable way\} is formed in a reliable way\} is formed in a reliable way,

he can, by exerting himself mentally, barely grasp it, although he also finds himself withholding it because of its complexity. You might think that this is exactly what one should expect to happen to rational people in a reflective mood drawn to continue considering propositions at ever higher levels: before reaching a proposition they are unable to grasp, they will reach one which they can barely grasp and which they will be inclined to withhold because of its complexity.

But as soon as that happens, the series of beliefs (from the highest higher-level belief in the series right down to the object-level belief that \( p \)) will come crashing down like a house of cards when a new card is clumsily added to it. For in withholding \( ((p^*_{Sam})^*_{Sam})^*_{Sam} \), Sam has an undercutting defeater for \( (p^*_{Sam})^*_{Sam} \), so he will, if he is rational, withhold it. But this gives him an undercutting defeater for his belief that \( p^*_{Sam} \), which rationally requires him to withhold this, which gives him a defeater for his belief that \( p \). So if what I have said above in the earlier sections is right, in order for rational and epistemologically sophisticated people to be justified in any of their
object-level beliefs, it seems that they must not get carried away with the sort of reflection just described.

(It is worth noting, in passing, that the ‘house of cards’ effect I have just described illustrates a non-contextualist way to account for something that contextualism accounts for. According to the contextualist, when we reflect critically on whether we know, we change the context in such a way that we can no longer truthfully say that we know. Likewise, when Sam reflects critically on whether he knows, it changes things in such a way that it is no longer correct to say of him that he knows. The difference is that according to the contextualist, the change in the truth-value of the knowledge attribution is due to the indexical nature of the term ‘knowledge’ and the way in which its meaning depends on the attributor’s context, whereas, in my ‘house of cards’ example, the change in truth-value is due to a change that reflection brings to the subject’s context by generating defeaters that were not present prior to reflection. Another difference is that according to the contextualist, the lack of knowledge after reflection is due to the belief’s failure to satisfy certain very high standards that are relevant given the new attributor context, whereas in the ‘house of cards’ example the lack of knowledge after reflection is due to the belief’s failure to satisfy an analytic requirement on justification, namely, that there must be no defeater for the belief.)

(b) A sceptical worry

But how does the house of cards effect create a problem for the externalist? After all, the point is just that people who engage in the sort of reflection just described will have defeaters for all their beliefs. Why can one not avoid the defeaters simply by avoiding the reflection? Is the problem not peculiar to the example? Is it not due to the fact that Sam happens to be helplessly addicted, at least on the occasion in question, to a certain kind of reflection?

Here is a way of using the house of cards effect to push towards scepticism externalists who accept the error theory given in §II, even if they do not actually engage in the sort of reflection described. A person who understands the house of cards effect will realize that there are at least two kinds of stopping-point in the process that leads to that effect: ceasing to reflect immediately after forming some higher-level belief $b$; or continuing to reflect until after one has considered the proposition $b$ is formed in a reliable way, but stopping before actually forming any doxastic attitude towards this proposition. Clearly the first kind of stopping-point is preferable. For at that point, one can, by taking no doxastic attitude at all towards the proposition

b is formed in a reliable way, easily avoid having a defeater for b. The reason why the second kind of stopping-point is not a desirable one is that it seems to force one into withholding the proposition b is formed in a reliable way. And, for the reasons noted above in §I(c), that seems to result in a defeater for b. These considerations will make anyone who is aware of the house of cards effect be careful to stop reflecting immediately after forming some higher-level belief b and before reflecting on whether b is formed in a reliable way. But that sort of active resistance to critical reflection on one’s beliefs seems like sticking one’s head in the sand. It does not seem like the way of true philosophy. For this reason, trying to pursue the sort of reflection in which Sam was helplessly engaged seems to be a rational thing to do, whereas actively resisting such reflection does not.

It seems that a sufficient condition for having a defeater for a belief is recognition that there is something rational one could do that is likely to give one a rationally acquired defeater for the belief. It is true that one could also do things that are likely to bring about irrationally acquired defeaters for a belief b (e.g., taking a delusion-inducing drug); recognizing this does not count as a defeater for b. Likewise, there are rational things one can do that might result in one’s having a defeater for a belief (e.g., drinking a glass of desperately needed water even though it might contain a delusion-inducing drug). Recognizing this does not give one a defeater for b either. But where one realizes that there is something rational one could or should do that is likely to result in one’s rationally acquiring a defeater for one’s beliefs, it is plausible to think that just realizing this constitutes a defeater. And, for the reasons given above, it seems that trying to reflect in the way Sam did is a plausible candidate for something rational one could or should do that is likely to result in rationally acquiring a defeater for one’s beliefs. It seems, therefore, that those who understand and accept the possible truth of my description of the house of cards effect, and who acknowledge that trying to engage in such reflection is a rational thing to do, will have a defeater for all of their beliefs, even if they do not actually engage in the reflection.

(c) Some responses

There are at least two types of response one could give to the above argument that accepting my account of the appeal of higher-level requirements forces one towards scepticism: (i) one could deny that reflecting in the way Sam does is likely to result in having a defeater for one’s beliefs; or (ii) one could deny that merely recognizing that Sam is likely to have a defeater for his beliefs after such reflection constitutes a defeater for one’s own beliefs.

There are at least two ways to develop response (i). You could say that S’s withholding p∗S does not always result in a defeater for S’s belief that p.
I said at the very end of §I that in the light of (D3*) and the discussion preceding it, it seemed that S would have a defeater for his belief that \( p \) if he withheld \( p^* \). But, strictly speaking, (D3*) says that S would have a defeater for his belief that \( p \) if significantly uncertain about \( p^* \), significantly enough to withhold \( p^* \). Thus, if S withholds \( p^* \) without being significantly uncertain about it, (D3*) does not say he has a defeater. And perhaps that is just what is happening to Sam. He reaches a point at which there is some higher-level proposition he barely grasps and on which he withholds judgement, not because he understands it and is uncertain about whether it should be believed, but because he is uncertain about whether he has grasped it clearly enough to be able to believe it. This is different from when people are quite confident that they have grasped a proposition and yet are uncertain about whether it is true. It may be, then, that this way of withholding \( p^* \) does not result in a defeater for \( p \), in which case the house of cards effect is prevented.

Another way to develop response (i) is to say that the attitude that Sam takes towards \( p^* \) is not withholding but merely the non-doxastic attitude of barely grasping. In order to resist believing a proposition or to resist disbelieving it (both of which are required for withholding it), one must grasp it clearly. If one barely grasps it, no doxastic attitude at all can be taken towards it. If this is right, then again the house of cards effect is prevented: there is no withholding, and therefore no defeaters (of the sort described).

As for response (ii), one could concede that Sam has a defeater for all the beliefs on which he reflects in the way mentioned, but insist that this is no reason to think that the rest of us who understand the effect have defeaters for our own beliefs when we do not engage in that sort of reflection. The strategy here would be to point out that trying to engage in that sort of reflection is not rational, and perhaps to note that if one does it compulsively, one has a lamentable (though minor and perhaps easily curable) cognitive defect. The suggestion is not that one should not reflect philosophically on epistemological issues, nor that one should not take doxastic attitudes towards higher-level propositions. Rather, the idea is that the sort of reflection in which one persistently continues, as Sam does, to take doxastic attitudes towards propositions at ever higher levels, until the propositions become too difficult to grasp, is not rational reflection, but is instead a symptom of a philosophical temperament gone awry. One can then think of the house of cards reflection as like the delusion-inducing drug: either of these may give one a defeater for one’s beliefs, but this is no reason to think that one’s beliefs are in fact defeated.

What I have shown in this paper is that by reflecting on the analytically true claim that our beliefs are not justified if we have a defeater for them, we can gain a better understanding of why it is so tempting to impose
higher-level requirements on justification (notwithstanding the fact that there are none). And although it appeared at first that this improved understanding forces one towards scepticism, I have shown that in the end it leaves intact externalism’s common sense assumptions about the justification of our ordinary beliefs.⁹

Purdue University, Indiana

⁹ Thanks to Jeffrey Brower, John Greco, Kevin Meeker, Trenton Merricks, Alvin Plantinga, Joel Pust, Michael Rea and several anonymous referees for comments on earlier drafts.