The No-Defeater Clause[†]

Evidentialism, Responsibilism and Higher-Order Evidence

— Penultimate Draft: please cite published version —

Forthcoming in *Episteme*

Simon Graf
Munich Center for Mathematical Philosophy, LMU Munich

Abstract:

Rational or epistemically justified beliefs are often said to be defeasible. That is, beliefs that have some otherwise justification conferring property can lose their epistemic status because they are defeated by some evidence possessed by the believer or due to some external facts about the believer's epistemic environment. Accordingly, many have argued that we need to add a so-called no defeater clause to any theory of epistemic justification. In this paper, I will survey various possible evidentialist as well as responsibilitst no-defeater clauses and develop a general taxonomy of defeater cases against which these clauses can be tested. Despite influential arguments that evidentialist understandings of justification are ill-equipped to handle the full spectrum of defeater cases, I will demonstrate that evidentialism has the right tools to make sense of all kinds of defeaters, including propositional and normative defeaters. Moreover, I will demonstrate that the proposed solution avoids recently influential objections against the notion of defeat.

1. Introduction

Rational or epistemically justified beliefs are often taken to be *defeasible*.[1] Take the case of perception. Many epistemologists argue that immediate perceptual beliefs provide us with justification in the absence of reasons to doubt them. For example, Pollock and Cruz state "if

[†] This article has benefited considerably from discussions, comments and feedback from Robbie Williams, Sandy Goldberg, Jessica Brown, Andrew Peet and Joshua Habgood-Coote, as well as anonymous comments from an associate editor and a referee from Episteme. The research leading to this article has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 818633).

[1] Note that I will discuss both factual defeaters, which are usually understood to be knowledge-defeaters, as well as justification-defeaters (Sudduth 2008; Graham & Lyons 2021). More on these distinctions in [§3].

something looks red to you and you have no reason to think that it is not red then you are permitted to believe it is red" (1999: 157). Similarly, Wedgwood thinks that you are permitted to believe that p when "you have an experience or apparent perception as of p's being the case, and have no special reason to think that your experiences are unreliable in the circumstances" (2002: 276). This gives rise to the notion of *prima facie* justification. Beliefs that have some initial justification-conferring property, such as being based on sufficient evidence or formed via a reliable process are *prima facie* justified.

This defeasibility of *prima facie* justification is a widely acknowledged feature of both internalist accounts of justification, such as evidentialism, as well as externalist accounts of justification, such as process reliabilism.[2] Furthermore, recent developments in collective epistemology suggest that understanding the role of defeat and its relation to group evidence are crucial elements in any analysis of collective justification.[3]

Let us call the doctrine that we can have *prima facie* justified beliefs which are defeasible, Defeatism:[4]

> **Defeatism:** Doxastic attitudes can have the status of being *prima facie* justified. That is, some doxastic attitude D can be *prima facie* justified by having some justification-conferring property P while being defeated and, therefore, lacking *ultima facie* justification.

Defeatism, while not without its critics, is motivated by a plethora of so-called defeater cases;[5] that is, cases in which it seems that some doxastic attitude D despite having the justification conferring

[2] For general discussions see Lehrer and Paxon (1969), Pollock (1986), Pollock and Cruz (1999), Lackey (1999), or Bergmann (2005). For externalist understandings of defeat see Goldman (1979), Alston (1988), Plantinga (1993: 40-42) or Graham and Lyons (2021).

[3] For notions of collective defeat see Schmitt (1994), Carter (2015), Lackey (2016, 2021), or Silva (2019). For related discussions of group evidence see Buchak and Pettit (2015) Hedden (2019) or Brown (2022).

[4] The term 'Defeatism' is taken from Baker-Hytch and Benton (2015).

[5] For critics of defeatism see, for example, Bergmann (2006), Lasonen-Aarnio (2014), or Baker-Hytch and Benton (2015).

property P is unjustified due to the existence of some defeater φ. Accordingly, many have argued that any theory of epistemic justification needs a so-called no-defeater clause, that specifies the conditions under which *prima facie* justified beliefs are *ultima facie* justified.

Usually, defeater clauses come in one of two forms:

> **Negative No-Defeater Clause:** S's *prima facie* justified belief that p is *ultima facie* justified iff there is no *defeater* φ (that S should have had), that is *defeating* the belief's *prima facie* justification.

> **Positive No-Defeater Clause:** If S's *prima facie* justified belief that p is *ultima facie* justified then S has a *defender* Φ, that is *defending* the belief that p from potential defeaters φ (that S should have had).

In this paper, I will survey various negative evidentialist as well as positive responsibilitst no-defeater clauses and develop a general taxonomy of defeater cases against which these clauses can be tested.[6] Despite influential arguments that evidentialist understandings of justification are ill-equipped to handle the full spectrum of defeater cases, I will demonstrate that evidentialism has the right tools to make sense of all kinds of defeaters, including propositional and normative defeaters.[7] Yet, I will do so in an unconventional way. Namely, I will identify two desiderata underlying responsibilist treatments of defeater cases and argue that we can understand these desiderata in light of a positive higher-order evidentialist clause. This is an important result, because, it allows those sympathetic to evidentialist understandings of epistemic justification to retain an evidentialist theory while taking into account defeater intuitions.

In particular, I will defend the following evidentialist higher-order clause:

[6] The term 'responsibilist' is taken from Cloos (2015).

[7] For discussions of normative defeaters see, e.g., Kornblith (1983), Lackey (1999), DeRose (2000), Baehr (2009), Cloos (2015), Goldberg (2016, 2017, 2018), or Graham and Lyons (2021).

Higher-Order Evidence Clause <HOE>:

S belief that p is *undefeated* iff:

(a) S *possesses* some higher-order evidence E_H, which is *sufficient to support* q, the proposition that the total relevant evidence E_{TOTAL} on balance supports p and

(b) S's belief that p is *properly based* on E_H.

Before I proceed, let me note that, while higher-order in some to-be-defined sense, <HOE> does not over-intellectualise justification in any problematic way as traditional internalist higher-order belief requirements arguably do.[8] Instead, as I will argue below [§8], <HOE> is cognitively very undemanding if we understand the underlying terms 'evidence', 'evidential support' and 'epistemic basing' in the right way.

Here is the outline. I will start with some conceptual remarks about evidence [§2] and various types of defeaters [§3]. These preliminaries enable me to illustrate how conventional evidentialist strategies fail to give us the right verdict concerning the full range of defeater cases [§4]. Analysing the ways in which these conventional evidentialist strategies fail, pushes us towards a responsibilist understanding of defeat [§5], and helps me to develop a general taxonomy of defeater cases [§6]. While this proposed taxonomy suggests a responsibilist solution, I will show that epistemic responsibility is best understood via the above-stated higher-order evidentialist condition [§7] [§8]. This not only sparks hope for those who want to defend an evidentialist understanding of epistemic justification in general but also allows us to put forward a universal treatment of defeater cases which could be understood as bypassing the doctrine of Defeatism [§9].

[8] For discussions of this worry see, e.g., Greco (1990), Bergmann (2005, 2006), (Lackey 2005) or Goldberg (2008).

2. Evidence

Before talking about defeaters, it is important to start with some preliminaries on evidence and its role in epistemic justification.

First, I will remain neutral concerning the *nature* of evidence. In other words, I will not take sides in the debate on whether pieces of evidence are mental states or propositions picked out by those mental states.[9] Relatedly, some argue that false propositions can be part of one's evidence, while those who defend a *factive* understanding of evidence claim that only true propositions can constitute evidence. Following Williamson (2000), factive understandings of evidence are increasingly popular. However, there is still a considerable number of epistemologists who understand evidence to be non-factive. Accordingly, I decided to organise the following discussions as if I relied on a non-factive understanding of evidence. Since if we can provide an evidentialist no-defeater clause relying on a non-factive understanding of evidence, we have done more than required for the evidentialists who have a factive understanding of evidence.

With these preliminaries about the nature of evidence in mind, we can now draw different distinctions between different types of evidence, such as *possessed* and *unpossessed*, *accessible* and *inaccessible*, or *available* and *unavailable* evidence.

Let us start with the evidential base, that is, the entire evidence *possessed* by an epistemic agent.

The evidential base (in short: E_{BASE}): The total evidence *possessed* by an epistemic agent S.

There are different ways to understand evidence possession. For example, mentalists might define E_{BASE} to be the entirety of some relevant mental states, such as experiences; whereas

[9] For a mentalist understanding of evidence see Conee and Feldman (2004). For further discussions see Turri (2009), McCain (2014: 10-11), or Sillins (2005).

perspectivalists hold that the evidence possessed by an epistemic agent is the evidence that they have a certain doxastic relationship with [§4].[10] This could, for example, be all of one's justified beliefs (if evidence is non-propositional) or all of the propositions one knows (if evidence is propositional).

Depending on our understanding of evidence and evidence possession, we can make some further distinctions here. For example, we may distinguish between the *accessible* possessed evidence and the *inaccessible* possessed evidence; where accessibility roughly means that S could become aware of the evidence upon reflection.[11]

The accessible base: The total evidence *possessed* by and *accessible* to S.

Based on this notion of accessibility we can further distinguish between accessible evidence that *has been accessed* (that is, evidence that S was or is aware of) and evidence that is accessible but *has not been accessed* (that is, evidence that S could become aware of).

The accessed base: The total evidence which is *accessible* to and *has been accessed* by S.

Furthermore, we could consider some normative restrictions on the accessible evidence to differentiate between the part of the accessible base which we *should have accessed* from the part which we could blamelessly ignore. Let us call this the required base:

The required base: The total *unaccessed* evidence that S *should have accessed*.

[10] For a discussion of perspectivalist constraints see, e.g., Alston (1988) and Greco (1990).

[11] Note that speaking of accessible evidence does not imply that the adequacy of the evidence is also accessible to the respective epistemic agent. In other words, we need to distinguish between what Alston calls "the accessibility of grounds" and "the accessibility of the adequacy of grounds" (1988: 276).

So far, we have introduced the notion of an evidential base and distinguished various subparts of it. However, beyond the evidential base is a wider set of evidence pertinent to epistemic considerations. It includes evidence that is not possessed by S but is epistemically relevant for S in a looser sense. It includes all the propositions that S should know and/or could know given some effort, or equivalently, all the evidence that S should come to possess and/or should come to possess. I will call this the *available evidence*.

> **The available evidence:** The total *unpossessed* evidence that S could *come to possess*.

The available evidence will be important in the following discussions of propositional and normative defeaters [§3]. One of the questions will be whether *prima facie* justified beliefs can be defeated only by unpossessed evidence that we *should have possessed*, or also by unpossessed evidence that we are not epistemically required to have but which is nonetheless available to S. That is, we could further differentiate between the unpossessed but required and unpossessed and not required evidence:

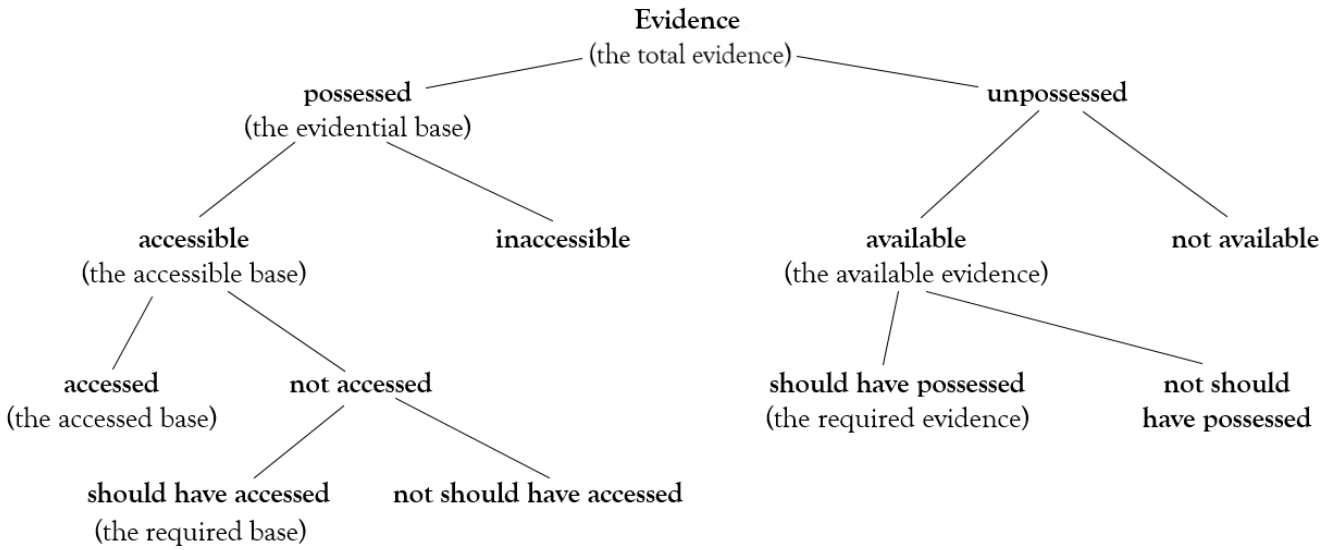
> **The required evidence:** The total *unpossessed* and *available* evidence that S *should have had*.

Finally, if we take the union of the evidence possessed by S and the evidence the evidence available to S we get what I call the total evidence:

> **The relevant evidence (in short: E_{TOTAL}):** The total evidence that is *relevant* to S's epistemic situation.

We can understand E_{TOTAL} as the entire evidence that bears on the proposition in question and is relevant to the epistemic situation. As such, it includes the entire evidence possessed by S as well as all of the evidence that is available to S in some important sense.[12]

In sum, we can distinguish between different subsets of the total evidence via the following descriptive relations *possession*, *availability*, and *accessibility*, as well as the two normative relations, *should have accessed* and *should have possessed*. This gives us the following tree diagram [Fig. 1].



[Fig. 1: Evidence]

Much of the following discussions will be centred around the questions of whether we can or should make all of these distinctions and whether some of these distinctions are extensionally equivalent. Regardless of how we answer these questions, having an intuitive understanding of these distinctions will turn out to be useful when characterising different kinds of defeaters [§3] and discussing extant no-defeater clauses [§4] [§5].

[12] At this point, we could also define an even larger set of evidence that includes all of the evidence, possessed and unpossessed, available or not, that bears on p. However, since E_{TOTAL} already includes all of the relevant evidence this larger set of evidence will be irrelevant to any of the cases discussed below.

3. Defeaters

Defeaters come in various types, and the two central kinds are *mental state* defeaters and *propositional* defeaters.[13] In other words, defeaters can be true propositions external to the perspective of the epistemic agent (propositional defeaters), or conditions internal to and/or within the perspective of the epistemic agent (mental state defeaters) (Sudduth 2008; §1).

Remaining neutral on the nature of evidence enables us to give a general evidential characterisation of defeaters that encompasses mental state defeaters, as well as propositional defeaters. That is, we can understand defeaters as evidence relevant to the epistemic situation of S that has some defeating force concerning some otherwise justified belief of S. Having this evidential understanding of defeat permits us to use the above-introduced relations (possessed, accessible, available…) to neatly distinguish various kinds of propositional and mental state defeaters.

First, *propositional defeaters* are outside of S's evidential base. That is, S's belief that p is factually defeated by some unpossessed evidence φ if acquiring φ would result in a loss of justification of S's belief that p (Sudduth 2008: §2; Bergman 2005: 154). For example, my otherwise justified belief that there is a barn in front of me might be defeated by the true proposition φ that I am in Fake Barn County (Goldman 1979). Accordingly, we can define propositional defeaters as follows:

> **Propositional Defeater:** Some evidence φ is a *propositional defeater* iff, (i) φ has enough defeating force to render S's *prima facie* justified belief that p unjustified and (ii), φ is *unpossessed* (i.e. outside E_{BASE}).

[13] For general discussions of the distinction between propositional and mental state defeaters see, e.g., BonJour (1980), Goldman (1986), Bergmann (2005), Lackey (2008) or Sudduth (2008).

These propositional defeaters are traditionally understood as defeaters for knowledge rather than justified beliefs. However, more recently, many epistemologists have argued that some specific kind of propositional defeaters, called *normative defeaters*, can also defeat or weaken the epistemic status of an otherwise justified belief.[14] If I should have known that I am in Fake Barn County, that is, if my ignorance with respect to φ is epistemically irresponsible (in a sense to be specified [§5]), φ is a normative defeater. Accordingly, we can characterise normative defeaters as "potential defeater[s] that the subject does not actually possess but should" (Graham & Lyons 2021: 45) or "evidence she does not possess but should have possessed" (Nottelmann 2021: 1183; see also Goldberg 2016; 2018: ch.6) or "a doubt or belief that S ought to have and that indicates that S's belief that p is either false or unreliably formed or sustained" (Lackey 2008: 45; see also Lackey 1999: 475; Lackey 2016: 366).

Accepting that there are normative defeaters raises further questions about the relationship between the evidence we should have possessed and its availability. In particular, we might ask ourselves whether 'ought have possessed' implies 'could have possessed' and/or whether 'could have possessed' implies 'ought have possessed'. While some have suggested that all normative defeaters are defeaters that are available (Harman 1980: 164; Goldberg 2018: 191; see also Nottelmann 2021: 1186), or at least indicated by the available or possessed evidence (Lackey 2016: 366), the literature is less clear on whether all available defeaters are normative defeaters. In acknowledgement of this, I will assume that all normative defeaters are available defeaters but not *vice versa*. An unopened letter containing some important information concerning the whereabouts of my friend might be an available, an unavailable, or an available and normative defeater for my belief that she is in town. For example, if the letter is on my desk it might be

[14] The term 'normative defeater' was coined by Lackey (1999). I will discuss her understanding of normative defeat in more detail below [§4.2.2].

available and normatively required, while it would only be available but not normatively required if it is at my colleague's desk and neither if it were still at the post office.[15]

Accordingly, we can define two types of propositional defeaters, available defeaters and normative defeaters, where the latter is a subtype of the former:

> **Available Defeaters:** Some evidence φ is an *available propositional defeater* for S's belief that p iff, (i) φ has enough defeating force to render S's *prima facie* justified belief that p unjustified and (ii), φ is *unpossessed*, *available* but *not required* (i.e. outside E_{BASE} but inside the required evidence).

> **Normative Defeater:** Some evidence φ is a *normative propositional defeater* for S's belief that p iff, (i) φ has enough defeating force to render S's *prima facie* justified belief that p unjustified and (ii), φ is *unpossessed*, *available* and *required* (i.e. within the required evidence).

Second, *mental state defeaters* are within the evidential base E_{BASE} of S. That is, a mental state defeater is some evidence possessed by S that has some putative defeating force with respect to some *prima facie* justified belief of S. Internalists characteristically deny that there are propositional or normative defeaters but both internalists and externalists typically agree that mental state defeaters can defeat otherwise justified beliefs.

> **Mental State Defeater:** Some evidence φ is a *mental state defeater* for S's belief that p iff, (i) φ has enough defeating force to render S's *prima facie* justified belief that p unjustified and (ii), φ is *possessed* (i.e. φ is within E_{BASE}).

[15] Compare Harman (1980).

Here, we could further distinguish between different kinds of mental state defeaters, including doxastic, reflective and inaccessible mental state defeaters.[16] A *doxastic defeater* is some piece of defeating evidence φ of which one is aware (φ is within the *accessed base*). A *reflective defeater* is some piece of defeating evidence φ of which one is not aware but of which one could become aware upon reflection (φ is within the *accessible base* but outside the *accessed base*). An *inaccessible* defeater is some piece of defeating evidence φ which is possessed by S but which S cannot become aware of upon reflection (φ is outside the *accessible base*).

> **Doxastic Defeater:** Some evidence φ is a *doxastic defeater* for S's belief that p iff, (i) φ has enough defeating force to render S's *prima facie* justified belief that p unjustified and (ii), φ is *possessed* and *accessed* (i.e. φ is within the accessed base).

> **Reflective Defeater:** Some evidence φ is a *reflective defeater* for S's belief that p iff, (i) φ has enough defeating force to render S's *prima facie* justified belief that p unjustified and (ii), φ is *possessed* and *accessible* but *not accessed* (i.e. φ is within the accessible base and outside the accessed base).

> **Inaccessible Defeater:** Some evidence φ is an inaccessible defeater for S's belief that p iff, (i) φ has enough defeating force to render S's *prima facie* justified belief that p unjustified and (ii), φ is *possessed* and *inaccessible* (i.e. φ is within E_{BASE} and outside the accessible base).

Whether these distinctions between doxastic, reflective and inaccessible defeaters are important, or whether they even make sense depends on various background assumptions about the nature of evidence and evidence possession, as well as assumptions about how defeaters exert their defeating force. For example, we could ask, following Alston (1988), whether accessible defeaters need to be accessible in the sense that we can access the evidence that constitutes the defeater or

[16] Similar distinctions are discussed by Bergmann (1998: 116-121) and Sudduth (2008: §5).

accessible in the sense that we can access the defeating force they exert. Without taking sides on this debate here, any answer to that question plausibly is related to general questions about evidential support. If we think that a piece of evidence can support (and also potentially justify based on that support) some proposition p without the awareness of the believer, we are likely to think that it can also exert its defeating power independent of whether that defeating power is accessible to the believer (more on that below [§8]).

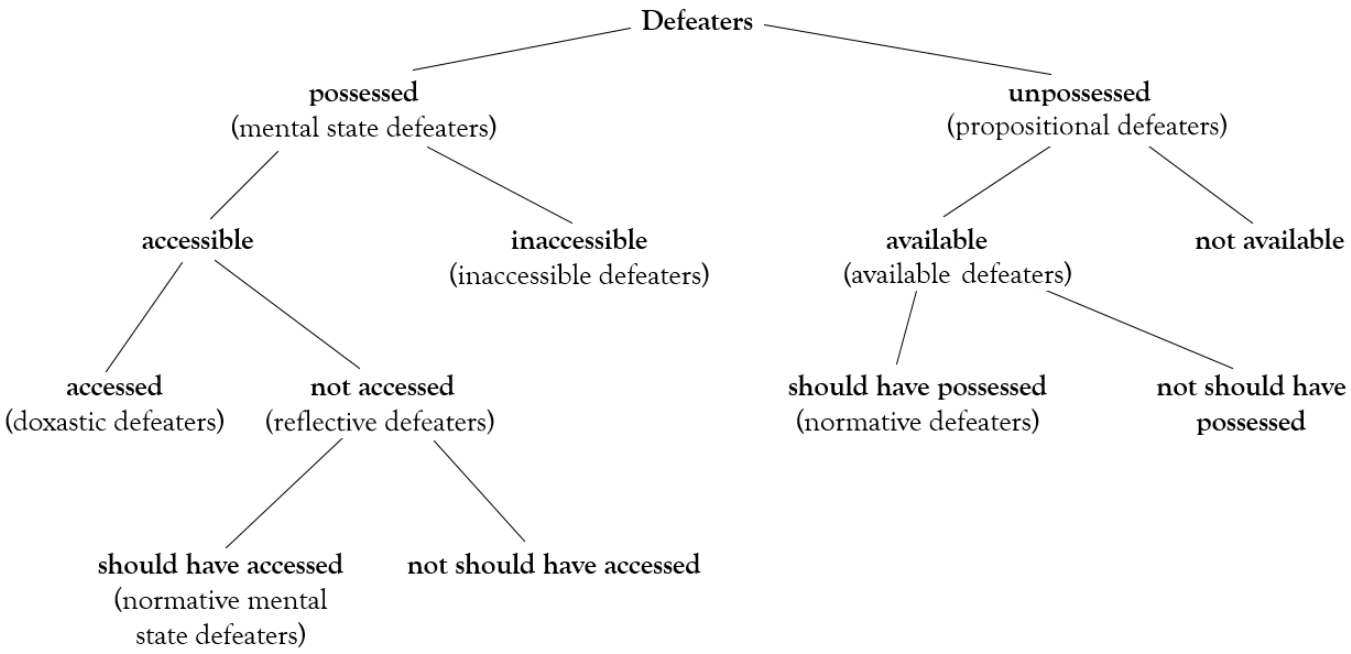
If we settle on the distinction between accessible and inaccessible mental state defeaters one way or another, one interesting question becomes whether there are defeaters which we possess and are not aware of but should be aware of. In other words, whether there is a second kind of normative defeater, which we may call *normative mental state defeater* or not:

> **Normative Mental State Defeater:** Some evidence φ is a *normative mental state* defeater for S's belief that p iff, (i) φ has enough defeating force to render S's *prima facie* justified belief that p unjustified and (ii), φ is *possessed*, *accessible* and *not accessed* but *should have been accessed*.

I think Jennifer Lackey, on various occasions, has convincingly argued that there are cases involving what I call normative mental state defeaters.[17] I will discuss one such case and take a closer look at Lackey's understanding of normative defeat in [§4.2.2].

In sum, this gives us two general types of defeaters, propositional and mental state defeaters, as well as various subtypes which inhabit different subparts of the entire evidence relevant to the epistemic situation E_{TOTAL}. Similar to how we characterised different subtypes of the relevant evidence above [Fig. 1] we can plug in different kinds of defeaters into a tree diagram [Fig 2]:

[17] See, for example, Lackey (1999: 75; 2006: 438-439; 2008: 45).



[Fig. 2: Defeaters]

Having categorised different types of defeaters we can now investigate how defeaters exert their defeating force.

First, we need to consider the question of whether defeaters need to have some positive epistemic standing. Does φ, to have a defeating force towards Bp, need to be sufficient to support B¬p or at least to repudiate the support that S had for p in the first place? Here the traditional answer is that a defeater is a reason or piece of evidence φ such that given the initial evidence E and φ together are not sufficient to support p (while E alone would be) (Pollock 1986; Pollock & Cruz 1999: 195; Graham & Lyons 2021: 40-41). Others have argued that mere beliefs without any positive epistemic standing can defeat otherwise justified beliefs (Plantinga 2000: 364-365; Lackey 1999, 2008: 44-45; 60-63). For Lackey, independently of whether doxastic defeaters are "true, justifiedly believed, rationally believed", what makes "an undefeated doxastic defeaters are epistemically problematic is that it is held in conjunction with another belief" (2005: 47). So, since contradicting unjustified beliefs can generate doxastic incoherence and *if* doxastic

incoherence amounts to defeat, both justified beliefs and unjustified beliefs can serve as defeaters.[18]

I am, following Graham and Lyons (2021), sceptical of the idea that the doxastic incoherence generated by unjustified beliefs can repudiate the justificatory status of well-supported beliefs. Because I think that the correct response to this doxastic incoherence is to drop the unjustified belief, rather than both beliefs or merely the justified one (Graham & Lyons 2021: 49-50). Nevertheless, I acknowledge that this is an ongoing debate and that there are interesting arguments to be made in favour of both views. Accordingly, the following investigations rest on two debated assumptions. First, I will assume that defeaters *need to have positive epistemic standing* to be able to exert their defeating force, and second, that it is evidence, not beliefs (unless those beliefs amount to evidence), which exerts defeating force. These assumptions will not only enable me to maintain the above-proposed analogy between different types of evidence and different types of defeaters [§2] – [§3], but also link the notion of defeating force directly to the notion of evidential support and the evidentialist higher-order clause I will end up defending in [§8].

Following Pollock (1986) I will assume that nothing can defeat a belief that cannot also provide justification. So, defeaters are either bodies of evidence strong enough to support justified beliefs or themselves justified beliefs. As such, there are different ways defeaters can exert their defeating force. Both mental state and factive defeaters can be either *rebutting*, that is, provide evidence that the belief that p is false, or *undercutting*, that is, provide evidence that the belief that p is unreliably formed or sustained (Pollock 1986). In Pollock's words, a rebutting defeater attacks the conclusion while an undercutting defeater "attacks the connection between the evidence and the conclusion, rather than attacking the conclusion itself" (1986: 38). For example, reliable testimonial evidence that my colleague is currently in France is a rebutting defeater to my belief

[18] For further discussion see, for example, Alston (2002), Bergman (2006: 164-166) or Graham and Lyons (2021: 47-52).

that I passed by her on campus today. In contrast, evidence about the unreliability of my facial recognition abilities, such as evidence that I have prosopagnosia or evidence about the unreliability of my short-term memory, is an undercutting defeater.[19]

Moreover, a defeater may itself be defeated (Lehrer & Paxson 1969: 228-229). In such circumstances, we may speak of a defeated defeater who is defeated by a *defeater-defeater*. For example, the allegedly reliable testimonial evidence that defeated my belief that my colleague is currently in France might itself be defeated by hearing from multiple independent sources that the person whose testimony I am relying on is a notorious liar. Differentiating between defeated and undefeated defeaters is important since it is often argued that while justified beliefs are incompatible with defeaters, we only need to worry about defeaters who have not been defeated themselves, that is, *undefeated defeaters*. (I will question this assumption below [§4]).

In sum, there are various kinds of defeaters including propositional (normative and non-normative) and mental state (including doxastic, reflective, inaccessible and normative) defeaters, which can defeat an otherwise justified belief in various ways (rebutting, undercutting, higher-order defeat), and which can itself be defeated. I will spend the next sections discussing conventional evidentialist [§4] and responsibilist [§5] strategies to account for the phenomenon of epistemic defeat. Based on these discussions, I will introduce a general taxonomy of defeater cases in [§6] which I will use to support my preferred understanding of defeat [§7] - [§8].

4. Evidentialism and Defeat

In this section, I will summarise conventional evidentialist ways of handling defeat and list some well-known, as well as some novel shortcomings of these approaches. Let me start with some preliminaries about evidentialism as understood by Conee and Feldman (2004).

[19] Finally, defeaters can exert their defeating force on a higher-order level, which arguably differs from these traditional means of defeat (Christensen 2010; Lasonen-Aarnio 2014).

The central evidentialist commitment in Conee and Feldman's (2004) framework is the following thesis:

> **Evidentialist Justification:** A doxastic attitude D towards proposition p is epistemically justified for S at t if and only if having D towards p fits the evidence S has at t.

Sometimes the core commitment of evidentialism is also expressed as a supervenience thesis, which states that normative facts about the doxastic attitudes of the epistemic agent directly supervene on facts about their evidence. That is, any two epistemic agents possessing exactly the same evidence would be exactly alike concerning what they are justified in believing about any given issue (Kelly 2016: §1; Fratantonio forthcoming: §1; Conee and Feldman 2004: 101).

However, many evidentialists take Evidentialist Justification as expressing only a necessary condition of epistemic justification rather than a sufficient condition, because even if the evidence possessed by an epistemic agent overall supports a given proposition, the way in which the respective attitude is formed can influence its normative status. So, if a belief is arrived at in an epistemically defective way, e.g., via wishful thinking, it is not fully justified even if it is supported by the possessed evidence. To this end, we need to distinguish between *propositional* and *doxastic* justification, where evidence alone only determines whether a belief is propositionally justified, for it to be doxastically justified, the belief also needs to be *properly based* on the evidence.[20] This gives us the following supervenience thesis:

> **Propositional Supervenience:** The propositional justification of anyone's doxastic attitude toward any proposition at any time strongly supervenes on the evidence that the person has at the time.

[20] Conee and Feldman (2004: ch. 4) speak of well-founded and non-well-founded justified beliefs. For an overview see Turri (2010).

So, how does this Conee and Feldman-style evidentialism do with respect to some simple mental state defeater cases?

> DEFEATER: S's belief that p is *prima facie* justified. However, S is not aware that she possesses a *reflective defeater* φ for her belief.

> DEFEATER-DEFEATER: S's belief that p is *prima facie* justified. However, S is not aware she possesses a *reflective defeater* φ for her belief. Luckily, φ is itself defeated by another *reflective defeater-defeater* λ.

The Conee and Feldman-style evidentialist seems to have a convenient way to account for mental state defeater cases such as DEFEATER and DEFEATER-DEFEATER. They can appeal to the notion of propositional justification. If the entire evidence possessed by an epistemic agent needs to support or fit the respective belief for the belief to be propositionally justified the belief cannot be defeated by a mental state defeater. Since propositional justification is usually understood to be a necessary requirement for doxastic justification, S's belief cannot be justified. So strictly speaking, for Conee and Feldman there are no mental state defeaters since there is no (propositional) *prima facie* justification if the belief is not supported by the entire evidential base of S.

Since we are looking for a no-defeater clause that can be added to all kinds of understandings of *prima facie* justification, we need to transform this insight into an independent no-defeater clause:

> **Evidentialist No-Defeater Clause <Evidentialism>**: S's *prima facie* justified belief that p is undefeated iff S's evidential base *on balance supports* p.

According to <Evidentialism> in DEFEATER-DEFEATER, S is justified in believing that p while in DEFEATER S's belief that p lacks justification. Despite this being the seemingly desired result, I will demonstrate that handling defeater cases, via appeal to propositional justification is mistaken. In particular, I will discuss two kinds of propositional defeater cases [§4.1] and two kinds of mental state defeater cases [§4.2] on which <Evidentialism> fails to deliver the right verdict.

4.1 Evidentialism and Propositional Defeaters

My first thesis is that, since <Evidentialism> is only concerned with defeaters within the evidential base, it cannot handle cases of propositional or normative defeat.

> PROPOSITIONAL DEFEATER: S's belief that p is *prima facie* justified. However, S is not aware that there is an unpossessed but *available defeater* φ for her belief that p.

> NORMATIVE DEFEATER: S's belief that p is *prima facie* justified. However, S is not aware that there is an unpossessed but *available defeater* φ for her belief that p that S *should have possessed*.

Depending on whether S should have possessed φ or not we have a template to construct normative propositional or merely propositional defeater cases. I will, for now, focus on normative defeaters. However, I will revisit this distinction at the end of the section. (Furthermore, the account I'll offer in [§7] can make sense of normative and non-normative propositional defeater cases).

The unique threat normative defeaters have for evidentialism has been pointed out by many, including Kornblith (1983), Baehr (2009), Cloos (2015), as well as Goldberg (2017: 2891-2893;

2018: ch.6) among others. If we are merely concerned with the overall evidence an epistemic agent possesses at a time, we can never consider the evidence (defeating or not) that is not possessed by S, but should be possessed by S. One illustrative example that has the structure of NORMATIVE DEFEATER is found in Kornblith (1983):

> PERSISTENT PETE: Pete a young and stubborn physicist presents his newest pet hypothesis at a conference eager to hear the praise of his colleagues. Due to a personality disorder, Pete pays no attention to reasonable critics and strategically ignores important counterevidence. As a result, even devastating criticism fails to impact his beliefs not because he fails to take it into account but because he has not even heard it.

For Kornblith this example illustrates that Pete's "belief is unjustified, after his colleague presents his objection, and it is unjustified because of his culpable ignorance" (Kornblith 1983: 36). As such, Kornblith offered the case as a direct challenge to evidentialism, since if his interpretation is correct there are non-evidential factors which determine whether Pete is justified or not.[21] While evidentialists have tried to dismiss the intuitive judgement we have towards cases such as PERSISTENT PETE, none of their strategies seems particularly convincing. For example, Conee and Feldmann (2004: ch.7) have argued that cases of normative defeat do not undermine epistemic justification but usually demark other shortcomings such as moral or professional failures. However, many disagree with this judgement (Goldberg 2016: 450; Lackey 2016: 374-375, Graham & Lyons 2021). After all, having these prudential or professional failures has devastating epistemic effects on Pete, not only concerning the reliability and truth-conduciveness of his belief-forming mechanisms but also concerning the evidence he possesses. There are good

[21] See also Goldberg (2016).

reasons to think that these allegedly professional failures are also epistemic failures (more on that below).

At this point, a plausible suggestion is to expand the evidence we are concerned with to include unpossessed evidence. That is, we might require that the total relevant evidence, possessed and unpossessed, on balance supports p.

> **Extended Evidentialist No-Defeater Clause <Extended Evidentialism>**: S's *prima facie* justified belief that p is undefeated iff the *total* evidence *on balance supports* p.

While <Extended Evidentialism> accounts for cases such as PERSISTENT PETE and other cases of normative defeat, it already demarks a significant departure from Conee and Feldman's version of internalist evidentialism, since it denies the central supervenience claims. Furthermore, it cannot account for the following type of normative defeater cases:

> LUCKY NORMATIVE DEFEATER-DEFEATER: S's belief that p is *prima facie* justified. However, S is not aware that there is an unpossessed but *available defeater* φ for her belief that p that S *should have possessed*. Luckily for S, there is another *normative defeater* λ which defeats φ.

To pump some intuitions, it might help to put some flesh on the bones:

> SORROWLESS SARAH: Sarah forms a *prima facie* justified belief that p. However, the belief is based on some information E drawn from the daily tabloid, unknowingly to Sarah a very unreliable source. This fact constitutes a normative defeater φ for the belief that p, since Sarah could have easily and should have learned about φ. Instead of checking the quality of her sources Sarah unreflectively formed her belief based on E. Yet, while generally unreliable the daily tabloid is reliable in this particular instance since the one

columnist Sarah based her belief on is exceptionally reliable. This fact serves as a defeater-defeater λ for φ.

Sarah's *prima facie* justified belief appears to be defective because she was just lucky that the defeater that she should have possessed is itself defeated. She could have easily learned about φ which, if rational, would have led her to abandon the belief. In not possessing φ Sarah behaved epistemically irresponsible just as Pete behaved epistemically irresponsibly when ignoring the counterevidence presented by his colleagues.

If that's the right verdict, we can easily see that extending <Evidentialism> to <Extended Evidentialism> does not solve the problem of normative defeat. Instead, normative defeater cases seem to suggest that we require epistemic agents to be responsible in the right kind of way. This is also illustrated by the following type of propositional defeater cases:

UNLUCKY PROPOSITIONAL DEFEATERS: S's belief that p is *prima facie* justified. However, while being generally a very responsible investigator S is not aware that there is an *unpossessed* and *available* defeater φ for the belief that p.

Here is such a case:

METICULOUS MICHAEL: Michael's belief that the tip of his tongue is mainly responsible for tasting sweet is *prima facie* justified. Michael has good evidence that this is true. He has once learned it in school and even remembers (apparently) confirming it himself as a kid. Furthermore, he even double-checks his belief by looking it up in his old biology book. However, while being generally a very responsible investigator Michael could have easily found out (e.g., via a quick Google search) that the tongue-tasting map is a common scientific misconception that has repeatedly been proven wrong. This fact serves as a propositional defeater φ for Michael's belief.

Can Michael's belief that p be justified despite there being a propositional defeater φ? It seems like it makes a difference if Michael acts meticulously or not. After all, he double-checked his belief and in general has good evidence for it being true. At least, there seems to be a difference to cases such as PERSISTENT PETE. Pete willfully ignored counterevidence and formed his belief partly based on the desire to prove his pet hypothesis.

If we judge Pete's and Michael's beliefs to have a different epistemic status, there are at least two ways to explain the different results. First, we might think that the way Pete and Michael form their beliefs makes those beliefs epistemically responsible or irresponsible respectively and hence justified or unjustified. On the other hand, we might think that the different judgements suggest that PERSISTENT PETE is a normative defeater case, where Pete should have possessed φ, while METICULOUS MICHAEL is merely a propositional defeater case, and hence Michael had no epistemic obligation to possess φ. Either way, the analysis goes through the notion of epistemic responsibility, which suggests a responsibilist treatment of defeat. I will explore this responsibilist treatment of defeaters further in [§5]. However, before plunging ahead, let me discuss another set of defeater cases evidentialist approaches have trouble dealing with [§4.2].

4.2 Evidentialism and Mental State Defeaters

While <Evidentialism> could handle regular mental state defeater cases, such as DEFEATER or DEFEATER-DEFEATER, normative defeater cases illustrated that <Evidentialism> is too narrow. In trying to overcome this problem by extending <Evidentialism> to <Extended Evidentialism>, we made it too general, since it declared sorrowless Sarah's belief to be justified despite her epistemically irresponsible behaviour and meticulous Michael's belief to be unjustified despite his epistemically responsible behaviour. We can find a similar pattern when we take a closer look at mental state defeater cases.

In particular, if we use propositional justification to declare all *prima facie* justified beliefs in all mental state defeater cases to be unjustified and in all defeater-defeater cases to be justified we get the wrong verdict with respect to two cases: *inaccessible defeater* cases, in which the epistemic agent is *responsible* [§4.2.1] and *accessible defeater-defeater* cases, in which a defeated defeater is *irresponsibly* ignored [§4.2.2]. Let me discuss those cases subsequently.

4.2.1 Inaccessible Defeaters. If we solely care about propositional justification, we rule out justification in *any* case in which we stipulate that there is a defeater φ within E_{BASE}. This is the case, irrespective of how strong the initial justification for Bp is (as long as φ has enough defeating force), regardless of whether the defeating evidence is doxastic, reflectively accessible or entirely inaccessible, and independently of how responsible the epistemic agent is. Here is a template for such cases:

> UNLUCKY MENTAL DEFEATERS: S's belief that p is *prima facie* justified. However, while being generally a very responsible investigator S is not aware that she *possesses* a non-doxastic *inaccessible* or *hardly accessible* defeater φ for the belief that p.

We can utilise this template to get the following case which bears some similarities to METICULOUS MICHAEL [§6.4.1]:

> CAREFUL CAROLINE: Caroline's belief that the tip of her tongue is mainly responsible for tasting sweet is *prima facie* justified. Caroline has good evidence that this is true. She has once learned it in school and even remembers (apparently) confirming it herself as a kid. Furthermore, she even double-checks her belief by looking it up in her old biology book. However, while being generally a very responsible investigator Caroline forgot that her university professor once told her that the tongue-tasting map is a common scientific

misconception that has repeatedly been proven wrong. This currently inaccessible memory serves as a non-doxastic defeater φ for Caroline's belief.

CAREFUL CAROLINE aims to illustrate that it depends on the exact details of mental state defeater cases if we should consider defeated beliefs to be justified or not. It seems overly demanding to require that epistemic agents are always completely sensitive to the entire body of evidence they possess, independently of how accessible the defeater is. After all, as justification is fallible concerning the truth of the formed beliefs it is also plausibly fallible with respect to the overall evidential support.

So, depending on the exact nature of the evidential base, <Evidentialism> may lead to an absurdly restrictive notion of justification. Take, for example, the so-called inclusive view of evidence discussed by Conee and Feldman (2004: 228). On this view, the evidence possessed by an epistemic agent is just the sum of all their non-factive mental states. That means the evidential base includes wrong beliefs and unconscious or in-principle inaccessible mental states. All of these mental states could serve as defeaters. In other words, even if S is overly responsible, attentive and aware of most of the evidence they possess the *prima facie* justified belief that p could ultimately be unjustified because they happen to possess an inaccessible or hardly accessible defeater.

Confronted with these cases the evidentialist might be tempted to restrict the notion of the evidential base in a way that rules out inaccessible or hardly accessible evidence. However, then the problem posed by normative defeaters becomes all the more important. In restricting the evidential base, we would push the defeater out of it into the unpossessed but available or required evidence and thereby transform cases such as CAREFUL CAROLINE into normative defeater cases such as PERSISTENT PETE.

4.2.2 Normative Mental State Defeaters: Furthermore, both <Evidentialism> and <Extended Evidentialism> also give us the wrong verdict for some normative mental state defeater-defeater cases. To see this, let us start with this case inspired by Lackey (2008: 63; see also 1999: 487):

> IGNORANT ALICE: Alice is told by her optometrist that her vision is nearly completely unreliable, yet she refuses to accept his diagnosis, without having any rational basis for doing so. Afterwards, as she is walking out of the doctor's office, she sees a car accident. Based on that Alice forms the corresponding true belief that there was such an accident. Intuitively, however, her belief is defeated by the doctor's diagnosis even if she refuses to accept it.

Let us compare this case to some cases already discussed. On the face of it, IGNORANT ALICE looks like a regular mental state defeater case rather than a normative defeater case such as PERSISTENT PETE. That is, ignorant Alice, in contrast to persistent Pete, possesses but ignores a defeater φ. Following Lackey, we could, nonetheless, call these kinds of defeater cases *normative* since Alice *should* form the corresponding belief. More precisely, Lackey thinks that a "normative defeater [...] function by virtue of being doubts or beliefs that S *should have* (whether or not S does have them) given the presence of certain available evidence" and further, that they exert their defeating force because "certain kinds of doubts and beliefs—either that one has or should have—contribute epistemically unacceptable *irrationality* to doxastic systems and, accordingly, knowledge (justification/warrant) can be defeated or undermined by their presence" (2008: 45; emphasis in the original). So, in this particular case, for Lackey, Alice should form the belief that her visual perception is unreliable because she possesses good evidence provided by the optometrist that would support this belief, whereas this belief would serve as a defeater, once formed.

While I agree with Lackey that IGNORANT ALICE constitutes a normative defeater case, I disagree about the way the testimony Alice received from her optometrist defeats Alice's belief about the accident. It is not the belief that Alice should have formed but the testimonial evidence

that serves as a defeater φ directly. This falls out of my view on how to understand defeating force [§3]. E (the witnessing of the car accident) together with φ (the testimony of the optometrist) do not together support p (that there was a car accident) and hence φ defeats Alice's belief that p.[22] This leaves us with the observation that we seem to have special normative obligations towards not only some evidence we *should possess* (as in normative defeater cases) but also some evidence we already possess but *should access*, which I call *normative mental state defeater cases*.

How does <Evidentialism> do with respect to normative mental state defeater cases, such as IGNORANT ALICE? At first, it seems like evidentialists can make perfect sense of them, since, after all, Alice's belief is *not* propositionally justified. Nonetheless, we can use the idea of having normative obligations to access some bits of our evidential base to construct the following type of defeater-defeater case which poses a problem for <Evidentialism>:

> LUCKY MENTAL DEFEATER-DEFEATER: S's belief that p is *prima facie* justified. However, due to being ignorant and negligent, S is not aware that she possesses an *accessible defeater* φ which she *should have accessed* for the belief that p. Luckily, S also possesses an *inaccessible* or *hardly accessible defeater-defeater* λ for φ.

We can again add some flesh to the bones:

> CARELESS CARL: Carl's belief that he saw his colleague Hao on campus today is *prima facie* justified. Carl has a vivid memory of her passing by the café while he was drinking his after-lunch coffee. Carl when asked if Hao is in town responds that she is without further reflection. However, another colleague told Carl last week that Hao will fly out on holiday tomorrow, something that Carl could recall easily if he would just reflect on his belief. This would make Carl doubt his memory and likely lead him to judge that he must have mistaken Hao for her twin sister. Accordingly, this memory counts as a defeater

[22] For a similar critique of Lackey's understanding of normative defeat see Graham and Lyons (2021: 52-56).

φ for Carl's belief that Hao is in town. Luckily for Carl, he also possesses a *defeater-defeater* λ for φ; namely that Hao told him that she did not plan any holiday this year and that she will definitely stay in town over the summer to work on her newest book. However, in contrast to φ, Carl cannot easily recall λ and hence would only remember it after a long period of reflection.

CARELESS CARL illustrates that not only in defeater cases but also in defeater-defeater cases the omitted details matter. Carl's belief, while undefeated and *prima facie* justified, seems epistemically defective for various reasons. First, Carl's belief is unstable. He could easily be aware of φ which, if rational, would lead him to dismiss the belief. So, it seems wrong to judge Carl's belief to be justified because of the luck involved. Carl, while having *prima facie* justified beliefs clearly forms his beliefs improperly. It is a mere coincidence that Carl's otherwise defeated belief ends up being undefeated. So, it seems like we have all the right to blame Carl for having this belief and generally forming beliefs in that way. And if this blameworthiness goes hand in hand with our epistemic evaluations, we need to say that Carl's belief is unjustified.

In sum, conventional evidentialist strategies fail to give us the right verdict in at least two different kinds of defeater cases. First, they are insensitive to propositional and normative defeaters, and second, they are too hasty in declaring all mental state defeater cases to be unjustified and all mental state defeater-defeater cases to be justified. While extending the pool of considered evidence (such as in <Extended Evidentialism>) helps with some propositional and normative defeater cases it does not help with the latter types of cases discussed.

5. Responsibilism: Two Desiderata

As demonstrated above, conventional evidentialism has problems with a plethora of defeater cases; namely, normative defeater cases (PERSISTENT PETE), normative defeater-defeater cases (SORROWLESS SARAH), lucky propositional defeater cases (METICULOUS MICHAEL)

inaccessible mental state defeater cases in which the epistemic agent is otherwise very responsible (CAREFUL CAROLINE) and defeater-defeater cases in which the agent behaves in an epistemically defective way (CARELESS CARL). On the face of it, these misjudgements of evidentialist strategies stem from a common source: they do not take into account whether the epistemic agents behave *epistemically responsible* or not.

Accordingly, one way to account for these cases is to introduce a no-defeater clause that obliges epistemic agents to be epistemically responsible in the right kind of way. On similar grounds, many have argued that we have epistemic normative obligations (Lackey 2016; 2021; Silva 2019), requirements to be sensitive to the evidence we should have had (Goldberg 2016; Cloos 2015) or that we are required to know everything we should have known due to professional or prudential duties (Goldberg 2017). Based on these insights Goldberg proposes an account of epistemic responsibility that is grounded in the epistemic expectations that epistemic agents are entitled to have towards each other (Goldberg 2018; ch.6). While I am sympathetic to this so-called entitlement model of epistemic responsibility, I will argue in [§8], that is entitlement is best understood as amounting to an entitlement to expect others to possess and utilize higher-order evidence that bears on their epistemic situation.

For now, however, we can lump these strategies together to specify a *responsibilist* treatment of defeat as follows:

> **Responsibilist No-Defeater Clause <Responsibilism>**: S's *prima facie* justified belief that p is undefeated iff S is *epistemically responsible* in believing that p.

<Responsibilism> looks promising, and indeed, if understood correctly, it gives us the right verdict with respect to all of the cases discussed so far [§6]. However, as I will demonstrate in [§7] and [§8], there are more and less plausible ways to understand epistemic responsibility. For now,

I will just, based on the discussions above, identify two desiderata any account of epistemic responsibility needs to fulfil:

> **Desideratum 1:** Any responsible epistemic agent forms and maintains doxastic attitudes in a way that is *sensitive* to the evidence they *accessed* and *should have accessed*.

> **Desideratum 2:** Any responsible epistemic agent forms and maintains doxastic attitudes in a way that is *sensitive* to the evidence they *should have possessed*.

These desiderata, while falling short of giving us a full-blown understanding of epistemic responsibility will be helpful when comparing different kinds of no-defeater clauses by checking their verdicts concerning the full taxonomy of defeater cases [§6].

Before moving on, however, let me briefly introduce a hybrid no-defeater clause that is neither exclusively evidentialist nor exclusively responsibilist.

> **Responsibilist-Evidentialist No-Defeater Clause <Responsible Evidentialism>**: S's *prima facie* justified belief that p is undefeated iff S's evidential base *on balance supports* p and S is epistemically *responsible* in believing that p.

Both Baehr (2009) and Cloos (2015) point out that there is no conflict, in principle, with combining responsibilist notions of defeat with evidentialist notions of justification.[23] However, these accounts are full-blown accounts of justification rather than merely no-defeater clauses that are used to supplement understandings of *prima facie* justification. In contrast, <Responsible Evidentialism> can be added to all kinds of understandings of *prima facie* justification including evidentialist, responsibilist or reliabilist accounts.

[23] For accounts of collective justification which have built in no-defeater clauses, and which combine responsibilist and evidentialist ideas see Lackey (2016: 381; 2021: ch. 2) or Silva (2019).

In sum, in [§3] - [§5], I have identified four responsibilist and/or evidentialist no-defeater clauses and discussed various defeater cases. The next section uses these insights to give us a general taxonomy of defeater cases against which any no-defeater clause can be tested.

6. Responsibility, Balance and Location: A Taxonomy

Most of the proposed no-defeater clauses make different judgements about the justificatory status of beliefs depending on the location of the defeater. For example, for <Evidentialism> and <Responsible Evidentialism> whether the belief that p is justified (in part) depends on whether the defeater is possessed or unpossessed (within or outside of E_{BASE}). <Extended Evidentialism> in contrast declares the belief to be defeated as long as there is a defeater within the total evidence E_{TOTAL}. Let us call the parameter that demarcates whether the defeater is possessed or unpossessed the *location* of the defeater.

In addition to the location, most evidentialist accounts care about the propositional justification, that is, the overall balance of the evidence. If there is a defeater φ (within E_{BASE}) the belief is unjustified; however, if there is a defeater-defeater λ that undermines φ the belief's justification is restored again. That is, the overall evidence can be in favour of p because there is no defeater, or because there are only defeated defeaters. Let us call this parameter the evidential *balance*.

<Responsibilism>, declares beliefs to be un/justified regardless of the overall balance and location of the defeater(s), while <Responsible Evidentialism> declares Careless Carl's belief to be unjustified and Careful Caroline's to be only justified if the defeater is *not* part of her evidential base. In contrast, <Evidentialism> and <Extended Evidentialism> do not take the careless or careful nature of the epistemic agent into account at all. Let us call this parameter the *responsibility* of the epistemic agent.

Having identified the parameters, *responsibility*, *balance*, and *location* presents us with eight possible cases, which, in turn, gives us a general taxonomy of defeater cases. Here is a template, we can use to change the parameters accordingly to see which account gives us which result:

> **Defeater Template:** S's belief that p is *prima facie* justified. There is a [*balance:* defeated/undefeated] and [*location:* possessed/unpossessed] defeater φ for S's belief that p. S is [*responsibility*: careless/careful] in forming and sustaining the belief that p.

In [Fig. 3] I plugged in the judgements made by <Evidentialism>, <Responsible Evidentialism>, <Extended Evidentialism>, and <Responsibilism> with respect to the different possibilities provided by Defeater Template, whereby 'N' stands for defeated or unjustified and 'Y' stands for justified or undefeated:

Taxonomy of cases:				**Evidentialism**	**Responsible Evidentialism**	**Extended Evidentialism**	**Responsibilism**
responsibility	**balance**	**location**	**verdict**				
careless	defeated	possessed	N	Y	N	Y	**N**
careless	undefeated	possessed	N	N	N	N	**N**
careless	defeated	unpossessed	N	Y	N	Y	**N**
careless	undefeated	unpossessed	N	Y	N	N	**N**
careful	defeated	possessed	Y	Y	Y	Y	**Y**
careful	undefeated	possessed	Y	N	N	N	**Y**
careful	defeated	unpossessed	Y	Y	Y	Y	**Y**
careful	undefeated	unpossessed	Y	Y	Y	N	**Y**

[Fig.3: Taxonomy]

By plugging in the respective parameters into the Defeater Template we can see that only <Responsibilism> matches our intuitive judgements: careless Carl's belief (careless; defeated; possessed) is not justified, such as persistent Pete's and sorrowless Sarah's (careless; defeated;

unpossessed) while Careful Caroline's belief (careful; defeated; possessed) remains justified [Fig. 3].

While this sounds like a full-blown responsibilist conclusion, the understanding of epistemic responsibility sketched in [§5] was quite schematic and hence is in need of further clarification. I will spend the next section [§7] adding futher details to the proposed understanding of epistemic responsibility, while I will argue in [§8] that the best way to understand epistemic responsibility is as a higher-order evidentialist requirement.

7. Responsibilism Revisited

In [§5] I have identified two desiderata underlying any responsibilist treatment of defeat that is able to give us the correct verdict with respect to the full range of defeater cases [§6].

While these two desiderata were helpful in specifying which shape any no-defeater clause should have, we lack a full account of epistemic responsibility. As a first step we could refer to the above-sketched understanding of epistemic responsibility as being grounded in socio-epistemic expectations (Goldberg 2016; 2017; 2018). This would allow us to make things more precise:

> **Epistemic Responsibility:** S is epistemically responsible if S is *sensitive* to (i) the evidence S *possesses*, including those which S has *accessed* and *should access* according to basic normative expectations, and *sensitive* to (ii) the evidence S *should possess* according to the epistemic expectations we are entitled to have towards S.

However, even if we accept Golberg's understanding of epistemic responsibility as being grounded in socio-epistemic expectations, a big question mark still hangs over the proposed understanding; namely, we still need to say what exactly it means for an epistemic agent to be *sensitive* towards defeaters in the right way.

Here, I think, as I have argued elsewhere (Graf 2024: 176-185), extant understandings of epistemic responsibility, such as the one proposed by Goldberg give us the wrong answer to this question. In short, Goldberg thinks that epistemic agents are strictly liable to the evidence they possess as well as the evidence they should have possessed (Goldberg 2018: 215), which gives us the wrong result with respect to some of the above-discussed cases. But instead of engaging in comparative analysis here, I want to put forward a positive proposal [§8].

8. A Higher-Order Evidence Clause

To recapitulate: we want a no-defeater clause that makes epistemic agents responsible, that is, sensitive to evidence, possessed or unpossessed, they should take into account, in a way that makes room for lucky and unlucky cases. In so doing, we want epistemically responsible agents to be less likely to ignore defeaters but rule in cases in which they are responsible but miss some defeaters due to no fault of their own. Furthermore, we want to rule out cases in which epistemic agents are irresponsible in forming some beliefs but are lucky that those beliefs are *not* defeated.

One promising way to retain these intuitive judgements about epistemic responsibility is to equate being epistemically responsible with *possessing* and *utilising higher-order evidence* about one's epistemic situation. That is, instead of taking epistemic agents to be liable for all of the evidence they have and should have had, we expect that epistemic agents have and utilise higher-order evidence about the entire evidence relevant to their epistemic situation E_{TOTAL}. So, we are not only entitled to expect people to not ignore their higher-order evidence about unpossessed and unaccessed evidence, but we are entitled to expect epistemic agents *to have such evidence in the first place*. This enables us to distinguish between lucky and irresponsible formed beliefs, such as careless Carl's beliefs, as well as unlucky and responsible formed beliefs, such as careful

Caroline's. We judge Carl to be irresponsible because he does not have and utilise higher-order evidence while we judge Caroline responsible because she does.

This suggests that being sensitive and hence epistemically responsible really just is possessing and utilizing higher-order evidence in the right kind of way. Let me further illustrate this idea with another example. For instance, we trust a doctor who makes a diagnosis based on some blood test result because of the doctor's expertise in reading and interpreting blood test results. This means that we are entitled to expect the doctor to have higher-order evidence supporting the general reliability of blood test results and his ability to interpret them. That is, we expect doctors to possess general information about blood tests via their medical education and specific information about their track record of making diagnoses based on blood tests. After all, if the doctor does not possess or use any such higher-order evidence in making their diagnosis we are entitled to blame them for that. So, for example, if the doctor has not had enough practice in making a diagnosis on blood test results, we expect them to indicate that. That is independent of whether the doctor's failure to use any such higher-order evidence leads him to misjudge the blood test results or not.

This indicates that we should *not* understand epistemic responsibility, while grounded in social expectations, via liability.[24] The doctor is *not* liable for not having and utilising evidence he should have in the case this leads to a misjudgement of the situation but is generally required to have higher-order evidence about her situation. Let us make this more precise. Let us stipulate that Dr. S looked at the blood test results E and concluded that p. Furthermore, Dr. S's belief is also based on a bunch of background evidence E_H about the general reliability of blood tests and her competence with respect to interpreting these test results. That is, E, as well as E_H played a non-deviant causal role in the belief formation of S's belief and/or S would provide or has the

[24] As suggested by Goldberg (2018).

disposition to provide E and E_H as reasons for p if asked (depending on the underlying understanding of epistemic basing).

In sum, this gives us the following picture of epistemic responsibility. First (i) running the risk of missing relevant evidence by not acquiring higher-order evidence about one's epistemic situation suffices for epistemic irresponsibility. Second, (ii) not running the risk but nonetheless missing evidence does *not* suffice for epistemic irresponsibility. Where responsibility is a necessary condition for having an undefeated and hence justified belief. This is the understanding of epistemic responsibility that underlies the higher-order evidentialist no-defeater clause announced above [§1]:

Higher-Order Evidence Clause <HOE>:

S belief that p is *undefeated* iff:

(a) S *possesses* some higher-order evidence E_H, which is *sufficient to support* q, the proposition that the total relevant evidence E_{TOTAL} on balance supports p and

(b) S's belief that p is *properly based* on E_H.

Let me now take a closer look at both subclauses of <HOE> and illustrate how they can be interpreted differently depending on different underlying understandings of *evidence*, *evidential support*, and *epistemic basing*.[25]

(a) states that epistemic agents need to have higher-order evidence E_H that is sufficient to support the proposition q, the total relevant evidence on balance supports p. So, one promising way to understand clause (a) is that epistemic agents need to have higher-order evidence in the form of a mental state (with the relevant propositional content) that supports q, which means sufficiently increases the probability of q, that E_{TOTAL} supports p.

[25] For an overview of different evidentialist treatments of these concepts see Fratantonio (forthcoming).

(b) encompasses the idea that merely stipulating that S possesses some second-order evidence E_H is not sufficient since this does not guarantee that E_H is involved in the formation of the belief that p. Without (b), E_H could not bear on the epistemic standing of Bp. So, as with beliefs formed via first-order evidence, we need a basing requirement. To illustrate this, we might think about careless Carl and careful Caroline again. If careful Caroline is responsible by having and utilising E_H it does not matter whether there is a defeater (possessed or unpossessed) for her *prima facie* justified belief. She is responsible because she tries to ensure that there are no defeaters. And, she did so by utilising sufficient higher-order evidence in her reasoning. In contrast, Carl, who forms beliefs on a whim, is irresponsible, even if he does coincidentally possess some higher-order evidence E_H. Subclause (b) encapsulates this intuition. It ensures that E_H is not merely possessed but actually used. That is, based on different understandings of epistemic basing, clause (b) could be understood in multiple ways, it could, e.g., mean that the belief that p is (non-deviantly) caused by E_H.

Note that given the above-outlined understandings of *evidence*, *evidential support*, and *epistemic basing*, accepting <HOE> as a necessary requirement on epistemic justification does not lead to any kind of *over-intellectualization*.[26] In other words, <HOE> is not very cognitively demanding, despite its higher-order nature. All it requires is that epistemic agents need to be capable of having two mental states with the relevant propositional content E and E_H that are sufficient to support (which, e.g. means increase the probability of) two different propositions and that E and E_H non-deviantly cause S to believe that p. None of this needs to be happening consciously, S could just be acquainted with a belief-formation procedure that utilizes higher-order and first-order evidence in the right way. This can happen without S being aware of E or E_H as well as the evidential support that E and E_H grant to the respective propositions p and q and without E and E_H and

[26] I want to thank an anonymous referee for useful comments on this issue.

their support being accessible to S.[27] This means that, in principle, not only adult human beings and other relatively cognitively sophisticated epistemic agents can form justified beliefs but also children and other (allegedly less sophisticated) non-human animals.

With these clarifications about (a) and (b) it is easy to demonstrate that <HOE> gives us the right verdict with respect to the entire taxonomy of defeater cases [§6]. Defeated beliefs can be justified independent of the overall evidential balance and the location of the defeater if the belief is partly based on higher-order evidence that sufficiently bears on the overall evidential situation. So, careless Carl and sorrowless Sarah have unjustified yet 'undefeated' beliefs, while careful Caroline and meticulous Michael have justified, yet 'defeated' beliefs. This is the case because one lacks while the other possesses and uses higher-order evidence that the total relevant evidence E_{TOTAL} supports the formed doxastic attitude. That the higher-order evidence is actually (properly) used is important since it suggests that the agent in consideration is in fact responsible and not only possesses sufficient evidence.

We can now take a step back and reflect on the initial idea of the defeasibility of justified beliefs which motivated our search for a no-defeater clause more generally and our analysis of epistemic responsibility in specific.

9. Concluding Thoughts

Our search for a no-defeater clause was motivated by the doctrine of Defeatism [§1]; which is the idea that doxastic attitudes can have the status of being *prima facie* justified (by having some justification conferring property) while being defeated. Defeatism is motivated by a plethora of defeater cases involving various types of defeaters. While I have demonstrated that responsibilist ideas motivate a specific understanding of defeat that can deal with the entire taxonomy of cases discussed [§6], I have also argued that the best way to spell out the underlying desiderata of

[27] In acknowledgement of Alston's (1988) distinction discussed in [§3].

responsibilism is as higher-order evidentialist requirement [§8]. This is good news for those who prefer an evidentialist approach to justification in general. Especially, since many have argued that evidentialism cannot make sense of many wide-shared intuitions with respect to defeater cases. Nonetheless, <HOE> is different from both extant evidentialist as well as responsibilist treatments of defeat.

First, <HOE> is different from the other evidentialist clauses discussed, since it is, like responsibilist requirements a *positive* rather than a negative no-defeater clause. As such, like other positive no-defeater clauses <HOE>, while motivated by defeater cases, does not necessarily commit oneself to the doctrine of Defeatism. After all, accepting <HOE> we do not need to distinguish between *prima facie* justified beliefs which can be defeated and *ultima facie* justified beliefs. Instead, higher-order evidential requirements could be understood as necessary conditions for justification and once a belief fulfils whatever the first-order conditions are plus the specified higher-order requirements the belief is justified, independently of whether there is a defeater or not (within or outside of E_{BASE}). This is in stark contrast to extant evidentialist understandings of defeat. These understandings rely on the notion of *prima facie* justification, which they spell out (partly) in terms of propositional justification [§4]. <HOE>, in contrast, is compatible with rejecting both of these ideas: we neither need to distinguish between *prima facie* and *ultima facie* justification, nor do we need to spell out *prima facie* justification in terms of propositional justification.

Second, <HOE> is importantly different from extant responsibilist no-defeater clauses. If we accept <HOE> as a general solution to defeater cases we are not committed to the idea that there are epistemic obligations that require us to possess certain bodies of evidence (as argued by many responsibilists). Given <HOE>, there is no evidence we *should* –epistemically speaking– have possessed or have accessed (except the higher-order evidence E_H that is required to be

responsible).[28] In that sense, most of the distinctions drawn in [§2] and [§3] are redundant. The only set of evidence we need to define to analyse the phenomenon of epistemic defeat is E_{TOTAL}, the entire evidence that is relevant to an epistemic situation, to judge whether some evidence E_H is sufficient to support that, E_{TOTAL} is sufficient to support p.

References:

Alston, W. (1988) An Internalist Externalism. *Synthese,* 74 (3), 265-83. https://doi.org/10.1007/bf00869630.

Alston, W. (2002) Plantinga, Naturalism, and Defeat. In James Beilby (ed), Naturalism Defeated? *Essays on Plantinga's Evolutionary Argument against Naturalism* (pp. 176-203). Cornell University Press.

Baehr, J. (2009). Evidentialism, vice, and virtue. *Philosophy and Phenomenological Research*, 78 (3), 545-567. https://doi.org/10.1111/j.1933-1592.2009.00255.x.

Baker-Hytch, M. & Benton, M.A. (2015). Defeatism Defeated. *Philosophical Perspectives*, 29, 40-66. https://doi.org/10.1111/phpe.12056

Bergmann, M. (2005). Defeaters and higher-level requirements. *Philosophical Quarterly* 55 (220), 419–436. https://doi.org/10.1111/j.0031-8094.2005.00408.x.

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

Brown, J. (2022). Group Evidence. *Philosophical Issues*, 32 (1), 164-179. https://doi.org/10.1111/phis.12220.

Buchak, L., & Pettit, P. (2015). Reasons and rationality: the case of group agents. In Hirose I. & Reisner A. (Eds.), *Weighing and Reasoning: Themes from the Philosophy of John Broome* (pp. 207–231). Oxford University Press.

Carter, J. A. (2015). Group Knowledge and Epistemic Defeat. *Ergo: An Open Access Journal of Philosophy,* 2. https://doi.org/10.3998/ergo.12405314.0002.028.

Christensen, D. (2010). Higher Order Evidence. *Philosophy and Phenomenological Research*, 81 (1), 185-215. https://doi.org/10.1111/j.1933-1592.2010.00366.x.

Cloos, C. (2015). Responsibilist Evidentialism. *Philosophical Studies*, 172 (11), 2999-3016. https://doi.org/10.1007/s11098-015-0454-9.

Conee, E. & Feldman, R. (2004). *Evidentialism: Essays in Epistemology*. Oxford University Press.

DeRose, K. (2000). Ought we to follow our evidence? *Philosophy and Phenomenological Research,* 60 (3), 697-706. https://doi.org/10.2307/2653824

[28] Note, that the requirement that we should have E_H is only a normative requirement to possess evidence in as much as all evidentialist theories require one to possess evidence that is sufficient to support our beliefs.

Fratantonio, G. (forthcoming). Evidentialism. In Kurt Sylvan (ed.) *The Blackwell Companion to Epistemology, 3rd edition*. Wiley-Blackwell.

Goldberg, S. (2016). On the Epistemic Significance of Evidence You Should Have Had. *Episteme,* 13 (4), 449-470. https://doi.org/10.1017/epi.2016.24.

Goldberg, S. (2017). Should have known. *Synthese*, 194 (8), 2863-2894. https://doi.org/10.1007/s11229-015-0662-z.

Goldberg, S. (2018). *To the Best of Our Knowledge: Social Expectations and Epistemic Normativity.* Oxford University Press.

Goldman, A. (1979). What is Justified Belief? In G. S. Pappas (ed.) *Justification and Knowledge* (pp. 1–23). Springer.

Goldman, A. (1986). *Epistemology and Cognition*. Harvard University Press.

Graham, P. & Lyons, J. (2021). The Structure of Defeat: Pollock's Evidentialism, Lackey's Framework, and Prospects for Reliabilism. In Brown J. & Simion M. (eds.), *Reasons, Justification, and Defeat* (pp. 39–68). Oxford University Press.

Graf, S. (2024). *On the Nature and Relationship of Individual and Collective Justification*. Dissertation. University of Leeds. https://etheses.whiterose.ac.uk/35226/.

Greco, J. (1990). Internalism and epistemically responsible belief. *Synthese*, 85 (2), 245 - 277. https://doi.org/10.1007/bf00484794.

Harman, G. (1980). Reasoning and Evidence One Does Not Possess. *Midwest Studies in Philosophy*, 5 (1), 163-182. https://doi.org/10.1111/j.1475-4975.1980.tb00403.x

Hedden, B. (2019). Reasons, Coherence, and Group Rationality. *Philosophical and Phenomenological Research*, 99, 581-604. https://doi.org/10.1111/phpr.12486.

Kelly, T. (2016). Evidence. *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/win2016/entries/evidence/.

Kornblith, H. (1983). Justified Belief and Epistemically Responsible Action. *Philosophical Review*, 92 (1), 33-48. https://doi.org/10.2307/2184520.

Lackey, J. (1999). Testimonial knowledge and transmission. *The Philosophical Quarterly*, 49 (197), 471–90. https://doi.org/10.1111/1467-9213.00154.

Lackey, Jennifer (2005). Memory as a generative epistemic source. *Philosophy and Phenomenological Research*, 70 (3), 636–658. https://doi.org/10.1111/j.1933-1592.2005.tb00418.x.

Lackey, J. (2006). Knowing from testimony. *Philosophy Compass*, 1 (5), 432–448. https://doi.org/10.1111/j.1747-9991.2006.00035.x.

Lackey, J. (2008). *Learning from Words: Testimony as a Source of Knowledge*. Oxford University Press.

Lackey, J. (2016). What is Justified Group Belief? *Philosophical Review,* 125 (3), 341–396. https://doi.org/10.1215/00318108-3516946.

Lackey, J. (2021). *The Epistemology of Groups*. Oxford University Press.

Lasonen-Aarnio, M. (2014). Higher-Order Evidence and the Limits of Defeat. *Philosophy and Phenomenological Research*, 88 (2), 314-345. https://doi.org/10.1111/phpr.12090.

Lehrer, K. & Paxson, T. (1969). Knowledge: Undefeated justified true belief. *Journal of Philosophy*, 66 (8), 225-237. https://doi.org/10.2307/2024435.

McCain, K. (2014). *Evidentialism and Epistemic Justification*. Routledge.

Nottelmann N. (2021). Against Normative Defeat. *Mind*, 130 (520), 1183–1204, https://doi.org/10.1093/mind/fzaa079.

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

Pollock, J. (1986). *Contemporary Theories of Knowledge*. Rowman & Littlefield.

Pollock, J., Cruz J. (1999). *Contemporary Theories of Knowledge*. Rowman & Littlefield.

Silins, N. (2005). Deception and Evidence. *Philosophical Perspectives*, 19 (1), 375-404. https://doi.org/10.1111/j.1520-8583.2005.00066.x.

Silva, P. (2019). Justified Group Belief is Evidentially Responsible Group Belief. *Episteme*, 16 (3), 262-281. https://doi.org/10.1017/epi.2018.5.

Sudduth, M. (2008). Defeaters in Epistemology. *The Internet Encyclopedia of Philosophy*. https://iep.utm.edu/defeaters-in-epistemology/.

Turri, J. (2009). The Ontology of Epistemic Reasons. *Noûs*, 43 (3), 490-512. https://doi.org/10.1111/j.1468-0068.2009.00715.x.

Turri, J. (2010b). On the Relationship between Propositional and Doxastic Justification. *Philosophy and Phenomenological Research*, 80 (2), 312-326. https://doi.org/10.1111/j.1933-1592.2010.00331.x.

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