

Defeaters

The concept of epistemic defeat or defeasibility has come to occupy an important place in contemporary epistemology, especially in relation to the closely allied concepts of justified belief, warrant, and knowledge (Shope, 1983; Pollock, 1986; Bergmann, 1997b, 2006; Steup, 1996). The latter are terms of positive epistemic appraisal or positive epistemic status. As a first approximation, *defeasibility* refers to a belief's liability to lose some positive epistemic status, or – if positive epistemic status is degreed – the liability of a belief to having this status downgraded in some particular way. For example, while *S* may be epistemically justified in believing some proposition *p* at time *t*, *S*'s belief that *p* might become less justified or unjustified at some later time *t**. Moreover, beliefs may also be prevented from having or acquiring some positive epistemic status in the first place. So more generally, defeasibility refers to a kind of epistemic liability or vulnerability, the potential of loss, reduction, or prevention of some positive epistemic status. A *defeater* is, broadly speaking, a condition that actualizes this potential.

TABLE OF CONTENTS

1. The Concept of Defeasibility
 - a. Defeasibility: Legal, Moral, and Epistemic
 - b. Defeaters in Epistemology: Basic Distinctions
2. The Gettier Problem and Propositional Defeaters
 - a. The Tripartite Definition of Knowledge and the Gettier Problem
 - b. Defeasibility Analyses and Propositional Defeaters
 - c. Constraints on Propositional Defeaters
3. Mental State Defeaters and General Epistemology
 - a. Internalism, Externalism, and Mental State Defeaters
 - b. Coherentism, Foundationalism, and Mental State Defeaters
4. Prominent Features of Mental State Defeaters
 - a. Newly Acquired State Defeaters and Newly Acquired Power Defeaters
 - b. Diachronic Aspects of Mental State Defeaters
 - c. Defeater-Defeaters
5. Variations on Mental State Defeaters
 - a. The Epistemic Status of Defeating Beliefs
 - b. Subjective and Objective Contours
 - c. Conscious and Reflective Defeaters
6. Taxonomy of Defeaters and Formalities of Defeat
 - a. Primary-Type Defeaters: Rebutting, Undercutting, and No Reason Defeaters
 - b. Secondary-Type Defeaters: Defeaters for Grounds of Inferential Beliefs
7. Conclusion
8. References and Further Readings

1. The Concept of Defeasibility

a. Defeasibility: Legal, Moral, and Epistemic

The language of defeasibility is not unique to epistemology. In fact, its use in epistemology is arguably derived from its use in legal and moral discourse. For example, H.L.A. Hart (1961) borrowed the term “defeasibility” from its prior uses in property interests and applied it to contracts. Hart explained that though contracts were comprised of an offer, acceptance and consideration, contracts may still be void or voidable due to some *exception* such as fraud or incapacity. In making this application to contracts, Hart noted that there is no specific term in the English language to refer to exceptions to a basic legal rule (Hart, 1961, p. 145; cf. Boonin, 1966). The defeasibility of legal rules is analogous to the defeasibility of moral rules in ethics or moral philosophy. While there may be obligations to do *X*, many ethical theories add that at least some of these obligations are only *prima facie* duties. They can be overridden by other factors and thus no longer morally binding. Moral rules, like legal rules, are subject to being defeated in particular circumstances or under particular conditions.

Talk of defeasibility in the legal and moral context translates into epistemic defeasibility in at least one obvious way. If we think of positive epistemic status as normative, then this status will - like moral and legal rules - be subject to being overridden by other factors. In circumstance *C* we may be epistemically justified to believe *p*, just as we are legally or morally justified to perform action *A* in circumstance *C*. In other circumstances *C**, though, we may no longer be epistemically justified to believe *p*, just as we are not legally or morally justified to perform action *A* in circumstance *C**. This is particularly evident in deontological conceptions of epistemic justification, according to which we have various intellectual obligations and certain epistemic principles forbid believing *p* under certain circumstances, for example when *p* is not likely to be true or when *p* is likely to be false. But even if we think of justification simply in terms of having adequate evidence, justification will be variable. Chisholm (1966, 1989, pp 52-69), for example, notes that while evidence *e* may make *h* evident, another evident proposition, *d*, may defeat the tendency of *e* to make *h* evident because the conjunction of *e* and *d* does not make *h* evident. In other words, there may be a loss of justification when new evidence is added to an existing evidence base.

b. Defeaters in Epistemology: Basic Distinctions

Defeater theories are generally distinguished by how they construe what does the defeating and what gets defeated.

(i) While some philosophers construe defeaters as conditions *external* to the perspective of the cognizer (true propositions), others construe them as conditions *internal* to the cognizer (mental states such as experiences or beliefs). Hence, while some philosophers might regard the true proposition <there is a blue light shining on the widgets> as a defeater for a belief about the color of the widgets, others would regard the subject's *belief* that <there is a blue light shining on the widgets> as the defeater. What does the defeating in the first case is a certain fact (the obtaining of which is independent of a

cognizer's beliefs or perspective). What does the defeating in the second case is a mental state of the cognizer.

(ii) Philosophers who construe defeaters as true propositions usually take defeaters to be conditions that prevent an overall justified true belief from counting as knowledge. So if the true proposition <there is a blue light shining on the widgets> is a defeater it would prevent my belief that <this widget is blue> from being something I *know*, even if this belief is justified and true. On the other hand, philosophers who take defeaters to be mental states of the cognizer tend to see them as defeating the justified status of a belief, either by downgrading the degree of justification or by canceling the justified status of the belief altogether. In this case, having a defeater for my belief that <this widget is blue> entails that this belief, even if true, is no longer justified or justified to the same degree. Of course, if justification (to some high degree) is necessary for knowledge, defeaters that defeat justification may also prevent a true belief from counting as knowledge.

2. The Gettier Problem and Propositional Defeaters

a. The Tripartite Definition of Knowledge and the Gettier Problem

One of the primary tasks of [epistemology](#) is the examination of the nature of *knowledge*. One aspect of such inquiry is the analysis of those conditions that are severally necessary and jointly sufficient for knowledge. There have been three fairly widespread and long-standing intuitions concerning knowledge in the Western philosophical tradition. First, a person *S*'s knowing some proposition *p* entails that *p* is true. Second, though more controversially, *S*'s knowing that *p* entails that *S* believes or assents to *p*, perhaps firmly. Third, knowledge is not equivalent to true belief. Knowledge has a certain surplus value over true belief. The ancient Greek philosopher Socrates indicated this surplus value metaphorically by speaking of knowledge as true belief that has been "tied down" or "tethered." Much of the work of epistemologists in the second-half of the twentieth century has been devoted to examining candidates for this epistemological tether, a plausible condition (or set of conditions) that can transform a true belief into knowledge. The term "justification" is commonly used to designate this condition. A justified belief is roughly one that has a positive tie or strong connection to the truth goal of believing, something like [evidence](#), grounds, reasons, or processes of belief formation that are in some sense indicative of the truth of the belief. The so-called traditional or tripartite definition of knowledge as *justified true belief* expresses all three of the above intuitions.

However, owing to Edmund Gettier's arguments (Gettier, 1963) epistemologists have generally recognized that justified true belief accounts of knowledge suffer from a basic defect or inadequacy. Gettier argued that there are cases in which an individual could plausibly be said to have a true belief that is justified but which fails to constitute knowledge. For example, I might be justified in believing that <either Jones owns a Ford or Brown is in Barcelona> because I validly deduce it from a justified belief <Jones owns a Ford>. If Jones does not own a Ford but Brown happens to be in Barcelona, I will have inferred a true justified belief from a false justified belief. However, it seems

counterintuitive in this case to suppose that I *know* that Brown is in Barcelona, even if the belief is true and justified.

One of the early proposals to handle the [Gettier Problem](#) involved adding a fourth condition to knowledge that excludes inferences from or dependence on any false beliefs (Shope, 1983, pp. 81-118). But Gettier cases can be generated where there is neither an inference from nor dependence on any false beliefs (Steup, 1996, pp. 15-16). So other strategies must be employed to deal with Gettier counterexamples. One of these strategies employs the concept of defeasibility or defeaters (Lehrer and Paxson, 1969; Swain, 1974; Shope, 1983;).

b. Defeasibility Analyses and Propositional Defeaters

Defeasibility analyses of knowledge come in a variety of different specific versions. The generic idea is that a person *S* knows *p* only if there is no true proposition, *d*, such that if *S* were to believe *d* (or *d* were added to *S*'s evidence for *p*), *S* would no longer be justified in believing *p*. In other words, the existence of certain *unpossessed* evidence prevents a person from actually knowing *p* if this unpossessed evidence *would* result in a loss of justification were the person to acquire the evidence, be aware of it, or recognize it. So according to defeasibility theories, it's a true proposition that does the defeating, not a believed proposition. Following Bergmann (2006, p. 154), I'll refer to these kinds of defeaters as *propositional defeaters*. So according to defeasibility analyses of knowledge we must adopt the view that:

[PD] *S* knows that *p* only if there is no propositional defeater *d* for *S*'s belief that *p*.

Consider the so-called "Fake Barn" scenario, an often-cited Gettier-type case used by Alvin Goldman (Goldman, 1976, pp. 772-73). Suppose Henry is driving through a Wisconsin town, admiring the scenery. He sees a barn and believes <there's a barn>. Unbeknownst to Henry, this Wisconsin town is full of papier-mâché barn facsimiles, which look like real barns when viewed from the road. However, the structure Henry happens to look at is a genuine barn. He just happens to glance in the direction where one of the few real barns is located. His belief is true since he's looking at a genuine barn. He also appears justified in holding this belief. Henry believes what seems to him to be the case. He has no reason to believe that anything is suspicious about his perceptions, much less that he's in a town mostly populated with fake barns. He also knows that barns are fairly common in this part of the state. Nonetheless, it seems that, however justified Henry may be in holding this belief, he doesn't *know* that there is a barn present. He is of course lucky to believe what is true in this circumstance, but it's precisely this feature of the situation that raises doubt about whether he knows there is a barn before him. Had he looked at any other time, his eyes would have landed on a fake barn and his resultant belief would have been false. Knowledge would seem to require that it not be a matter of epistemic serendipity that one's belief is true.

Defeasibility analyses of knowledge attempt to relate the problem of accidentally true belief to the existence of some bit of relevant unpossessed evidence. That is, it is in

consequence of lacking some relevant evidence, of being less than ideally situated with respect to the evidence, that a person ends up luckily believing what is true. This is illustrated in the Fake Barn scenario. In that case, there is a true proposition such that, if Henry were to believe it, he would not have been justified in believing that the object he sees is a barn. The true proposition would be something to the effect that <in this town nearly everything that looks like a barn isn't actually a barn>. Call this proposition *D*, and call his barn belief *B*. If Henry were to believe *D*, he would not be justified in his belief that *B*. Alternatively, we might say that if *D* were added to Henry's actual evidence *E* (the evidence of his senses and relevant background beliefs), he would no longer be justified in holding the belief that *B*. Given *E*, Henry is justified to believe *B*, but given the conjunction of *E* and *D*, Henry is not justified in believing *B*. For Henry to know that there's a barn present, it must not be an accident that this belief is true. This in turn requires that Henry's justification be indefeasible.

We should underscore that there being a propositional defeater for Henry's belief that <there's a barn> does not entail that Henry is actually unjustified in believing <there's a barn> or that he's irrational or unreasonable in holding this belief. The point about justification is a counterfactual one: Henry *would not* be justified in believing <there's a barn> if he *were* to believe <in this town nearly everything that looks like a barn isn't actually a barn> or if this fact were added to his evidence. The counterfactual truth about justification entails that Henry doesn't actually know <there's a barn>, not that he's unjustified in believing it. Of course, if we're thinking of knowledge as simply justified true belief, we might speak of Henry's *justification* being defeated in some way because the justification is insufficient for knowledge (Lehrer and Paxson, 1969). The target belief may be justified, but the justification is "defective" (Marshall Swain, 1981, p. 148) because it fails to make his true belief knowledge. Steup (1996, p. 15) captures this point by speaking of the epistemizing potential of a person's justification being defeated, and contrasts this with saying that a person's justification is defeated. While Shope (1983, p. 47) speaks of *S*'s actual justification being defeated, by this he simply means that the justification fails to be enough – together with the satisfaction of the truth and belief conditions – for knowledge. And so also with other authors who use similar language at this juncture. So we should say that a propositional defeater for *S*'s belief that *p* doesn't entail that *S* is no longer justified in believing *p*, only that *S*'s justification isn't sufficient (along with true belief) for knowledge. Technically, then, we should speak of *knowledge* being defeated (Audi, 1993, pp. 185-213) or *warrant* being defeated (Plantinga, 2000, p. 359-60), where warrant is the property that transforms true belief into knowledge.

c. Constraints on Propositional Defeaters

As widely discussed in the early literature on defeasibility theory (Lehrer and Paxson, 1969; Annis, 1973; Swain, 1974), the main challenge facing defeasibility analyses of knowledge is to specify the relevant range of true propositions that can function as defeaters. It is generally acknowledged that not just *any* true proposition (suggestive of a defect in justification) is an efficacious defeater. There are genuine defeaters, but there are also misleading defeaters.

In the famous so-called Tom Grabit case (Paxson and Lehrer, 1969), I see a man who looks to me like Tom Grabit remove a book from a library bookshelf, slip it under his coat, and escape the library. I believe that Tom Grabit stole a library book. As it happens, the man I saw was indeed Tom Grabit, and he did steal the book. However, let's suppose further that Tom Grabit's mother claims that on the day in question Tom was out of the country but that Tom's identical twin brother John was at the library. Here it seems that there is a true proposition such that if I were to believe it, I would not be justified in believing that Tom Grabit stole a library book. The true proposition is <Tom Grabit's mother is testifying that. . . >. Call this true proposition *D*, the ostensible defeater. It would seem that, like in the case of Fake Barn, there is a propositional defeater for the target belief. I may in fact have a justified true belief that there is a barn over there, but the justification is defective and so my justified true belief does not constitute knowledge. The true proposition *D* is such that if I were to believe it (or add it to my evidence), I would no longer be justified in believing that Tom Grabit stole the library book. But now suppose that Mrs. Grabit is actually a compulsive liar and Tom's twin brother is the product of Mrs. Grabit's demented imagination. Tom Grabit is not out of the country and he has no twin brother. Given this expansion of facts, our intuition may now be that I *do* know that Tom Grabit stole a library book, that Mrs. Grabit's testimony does not actually defeat my knowing that Tom Grabit stole the book.

While we might say that there is a propositional defeater for my belief that Tom Grabit stole the library book, we can say one of two possible things about the defeater's lack of defeating efficacy.

First, the defeater in the Tom Grabit case is clearly misleading. It is perhaps natural to say that it misleadingly suggests that that the target belief is false or that the evidence for the target belief isn't good. The defeater is a true proposition, for it is true that Mrs. Grabit *said* that Tom's twin brother, not Tom, is in the library, and that Tom is out of the country. The problem is that this true proposition suggests that my belief that Tom Grabit stole the book is false or that I shouldn't be relying on the evidence of my senses. It also suggests other false propositions, for example that Tom Grabit has an identical twin, that Tom was not at the library, or depends on the false assumption that Mrs. Grabit is sane and her testimony reliable. At all events, what is required is a genuine as opposed to misleading defeater, and such a defeater will not presuppose, suggest, or depend upon some falsehood (Klein, 1976, 1981).

Secondly, we might say that the potential defeating effect of *D* is neutralized or defeated by some further true proposition, *D**, such that if I were to believe *D** I would not be justified in believing *D*. In this case, the true proposition, *D**, is that Mrs. Grabit is a liar and mentally deranged, whereas *D* is simply the fact of her testimony. It seems that *D* defeats my belief that Tom stole the library book because if I believed *D*, I would cease to be justified in believing he stole the book. But if I were to believe *D**, I would not be justified in believing the content of Mrs. Grabit's testimony. In other words, the total evidence set includes *D* and *D**, but *D** defeats *D*. A genuine defeater must be undefeated by any further evidence (Barker, 1976; Pollock, 1986; Swain, 1974).

Other epistemologists suppose that what defeats knowledge is unpossessed evidence that most of the members of the person's society or social group are aware of. We can use the example provided by Gilbert Harman (1973, pp. 143-44). Suppose that a political leader has been assassinated. A reporter who is a witness to the assassination dictates details of the event to his news agency so that the story may be included in the day's final edition of the paper. Jill picks up the paper and reads the story and believes that the political leader has been assassinated. However, before Jill picks up the newspaper and reads the story, loyalists to the political leader declare on nationwide television that the bullet actually struck and killed someone in the political leader's entourage. Jill reads the true story in the paper but misses the false report on television. Harman contends that in this hypothetical situation Jill doesn't know that the political leader has been assassinated. Some epistemologists (Swinburne, 2001; Pollock 1986) contend that Jill's not having knowledge in this case is the consequence of there being a true proposition (suggestive of a defect in justification) that is widely believed in Jill's society. (Advocates of this view would also seem committed to saying that if the Tom Grabit example were altered so that Mrs. Grabit testified in a public venue to the alleged whereabouts of Tom and the existence of Tom's identical twin brother, then her testimony would be a genuine defeater for someone's knowing that Tom stole the book, even if Mrs. Grabit were lying or deranged).

Alternatively, we might suppose that the crucial factor that determines whether a true proposition (suggestive of a defect in justification) is an efficacious defeater is if the unpossessed evidence is the sort of thing that is easily accessible. We can take another example from Harman (1973). Suppose your good friend Donald tells you that he's going to Italy for the summer. You take him to the airport and see him off. He left in June, but in July he decides to send you several letters informing you that he's actually in San Francisco. This is not true. He's simply trying to fool you. He sends the letters to another friend in San Francisco who is instructed to send them to you one at a time, as if they were sent from Donald, complete with a San Francisco postmark. You've been gone for a couple of days, though, and your mail has piled up. There are two letters in the stack from Donald. You haven't looked at them yet and so you believe that Donald is in Italy. This is true, but there's evidence of which you are not aware that would justify you in believing that Donald is not in Italy. It might be argued that in this case, the information contained in the unopened letters constitutes a genuine defeater for your belief that Donald is in Italy since the information is near at hand, readily available to you, even though in fact you're not aware of it.

There are of course other variations on genuine defeaters. We might throw a deontological spin into the defeasibility account. We might suppose that unpossessed evidence defeats knowledge only if the evidence is the sort of thing the person should believe and would believe if certain intellectual obligations were satisfied. At all events, all these defeasibility formulations are ways of placing constraints on propositional defeaters. They each recognize that while there are many true propositions that seem to indicate a defect in justification (that is, such that if S were to believe them, S would cease to be justified in his original belief) only some of these entail an actual defect in one's justification, actually defeat the person's knowing the target proposition.

3. Mental State Defeaters and General Epistemology

While defeasibility accounts of knowledge take defeaters to be facts external to the perspective of the cognizer, another approach to defeaters construes them as items internal to the perspective of the cognizer, as mental states such as experiences, beliefs, or withholdings. For example, on a particular day I see a person who looks like Tom Grabit steal a book from the library. Based on my sensory perceptual experience and my memory beliefs about what Tom Grabit looks like, I believe that Tom Grabit stole the library book. Later that day Tom Grabit's mother tells me that Tom is out of town but that his kleptomaniac identical twin was at the library at the time in question. Unlike the case of propositional defeaters, the defeater here is information I actually possess, something I learn or come to believe. It may not even matter that Mrs. Grabit is in fact a liar or delusional, unless of course I have reason to believe that this is true. Following Bergmann (2006, pp. 154-55), I'll refer to these kinds of defeaters as *mental state defeaters*. (Some philosophers, for example Alston 1986, p. 191, refer to these as "overrides" and reserve the term "defeater" for propositional defeaters. This terminological point is worth noting, but nothing substantive rides on this).

Epistemologies that incorporate mental state defeaters typically take them to defeat justification (Alston 1989, pp. 238-39; Bergmann, 2006, pp. 155-56) or some species of rationality (Plantinga, 2000, pp. 357-66; Bergmann, 1997a, pp. 68-78). However since these positive epistemic statuses are typically regarded as necessary for knowledge, mental state defeaters may at least indirectly play a role in defeating knowledge, not simply by *preventing* a person from coming to know *p* but also by *canceling* a person's state of actually knowing *p*. If *S*'s knowing that *p* entails that *S*'s is justified to degree *N* in believing *p*, then if *S* ceases to be justified in believing *p* (or the degree of justification for *S*'s belief is significantly lowered), then *S* ceases to know *p*. So we can think of mental state defeaters as defeating one's actual justification and knowledge. We can refer in a general way to a *no mental state defeater condition* for knowledge:

[MSD] *S* knows that *p* only if *S* does not have a mental state defeater for *S*'s belief that *p*.

Note that [MSD] only claims that knowledge requires the absence of a mental state defeater, a defeater constituted by a person's experience(s), belief(s), or other propositional attitudes. It doesn't specify or delimit the range of *what* mental states will actually count as defeaters. Would, for example, my simply taking a belief to be defeated count as a mental state defeater? Or must I justifiably take a belief to be defeated? Or must there be some kind of logical relation between my beliefs and the defeatee? Similarly, must mental state defeaters be occurrent states or can they be merely dispositional? Advocates of [MSD] disagree about these issues, as we'll see below. But the general idea behind mental state defeaters is a fairly bipartisan epistemological insight, as may be shown by its place in the broader landscape of contemporary epistemology.

a. Internalism, Externalism, and Mental State Defeaters

Epistemic [internalists](#) typically recognize that mental state defeaters can defeat justification (Pollock, 1974, 1984, pp. 200-202, 1986, pp. 29-30, 37-58; Chisholm, 1989, pp. 55-60; Swinburne, 2001, pp. 28-31). For the internalist, the endorsement of [MSD] is largely a consequence of justification supervening solely on the perspective of the cognizer. Just as the subject's beliefs and experience confer justification on beliefs, they can also remove or downgrade justification. If we also suppose that justification is necessary for knowledge, the internalist will endorse a principle similar to [MSD]. Of course, for the internalist [MSD] is not an alternative to [PD]. [MSD] doesn't address the Gettier problem but only concerns evidentialist intuitions about justification. [PD] is still needed by internalists to handle Gettier cases. But note also that the explication of [PD] seems to depend on certain counterfactual claims about mental state defeaters and justification, for we must suppose that if *S* were to believe *d* (or we were to add *d* to *S*'s evidence), then *S* would no longer be justified in believing *p*. This presupposes that one's actual evidence can defeat one's justification. In this way [PD] presupposes the type of conceptual framework employed by [MSD].

Many [externalists](#) have endorsed [MSD]. For example, some reliabilists (Goldman, 1986, pp. 62-63, 111-112) include a non-undermining provision in their accounts of justification or knowledge. In consequence of such a provision, while reliability of belief formation may be a necessary condition for knowledge, it's also necessary that a person not (justifiably) believe that his belief was formed in an unreliable manner. Alston (1988a, pp. 238-239) contends that truth-conducive justification can be *overridden* by justified beliefs that *p* is false or the justified belief that the belief that *p* is based on inadequate grounds. According to Plantinga (1993a, pp. 40-42, 229-37; 2000, pp. 359-66), while warrant depends on the proper functioning of our truth-aimed cognitive faculties, one aspect of this proper functioning is a sub-system (called a defeater system) that adjusts or revises our beliefs in the light of new experiences and beliefs. Nozick (1981, p. 196) contends that knowledge requires that the subject not believe that her belief doesn't track truth. In each of these cases, the otherwise externalist theory advocates at least one internal condition for knowledge, roughly that the subject does not have a negative epistemic evaluation of her beliefs.

b. Coherentism, Foundationalism, and Mental State Defeaters

The idea that mental state defeaters can cause justified beliefs to become unjustified (and the correlated [MSD] condition) is compatible with [coherentism](#) and foundationalism, and is arguably entailed by some versions of each.

From a coherentist viewpoint, coherence (of some form) among our mental states confers justification on our beliefs. Very roughly stated, I am justified in believing *A* if and only if *A* coheres with my current experience and body of beliefs. It follows that I will *become* unjustified in holding some belief *A* if the belief *A* loses its coherence with my experience or body of beliefs. But a belief's losing coherence with our experience or beliefs is a particular way of unpacking the idea of mental state defeaters. For example, I might at time *t* recall the foyer of a certain Victorian house in Springfield, Massachusetts having

certain structural features, and there's no incoherence at time t between my beliefs about the foyer and the rest of my experience or beliefs. However, upon subsequently revisiting the house at time t^* I see that it's not at all as I remember it. My present sensory experience is incompatible with my memory beliefs about the foyer and so my former beliefs about the foyer now become unjustified. Upon being appeared to catly, I may believe that there is a cat in front of me. This belief may cohere with everything else I believe and am currently experiencing at the time, so it's a justified belief. But suppose that when I reach out for the cat my hand goes through it, or when I move a couple of feet to the right or left the cat disappears and then reappears when I move back into place. My belief that there's a cat in front of me no longer coheres with the larger network of my beliefs. In this scenario I have lost my justification for supposing that there's a cat in front of me.

Mental state defeaters also play an important role in many versions of foundationalism, specifically versions of so-called modest foundationalism (Alston, 1976, 1983; Audi, 1993). Foundationalist theories of justification, motivated largely by the justification regress problem, terminate chains of justification in foundational beliefs that are immediately justified. Immediately justified beliefs are beliefs that are justified in some way other than their relation to or dependence on other justified beliefs. Strong versions of foundationalism restrict foundational beliefs to beliefs with various epistemic immunities (from doubt, error, or revision) or beliefs that are ostensibly maximally justified. These versions of foundationalism have little or no place for the idea that subsequent mental states might cause immediately justified beliefs to become unjustified (or less justified). But this idea is important to modest foundationalists, who argue that the regress problem may be avoided if chains of justification terminate in beliefs that are *prima facie* immediately justified. I can be immediately justified in believing that there is a cat in front of me, even if I subsequently lose this justification by realizing that I'm looking at a papier-mâché cat. My justification is in the first instance *prima facie* and thus capable of being overridden, cancelled, nullified, or downgraded by new experiences or additions to my beliefs.

Audi (1993, pp. 105-112, 141-53) notes that one of the core intuitions behind coherentism is really the idea of "negative epistemic dependence," that a belief's justification is liable to being overridden or undermined and so should not remain unaffected by incoherence if it should arise. A belief that is justified at time t independent of its relation to other beliefs need not be such that it remains justified (or justified to the same degree) regardless of the other beliefs a person forms. The idea of mental state defeaters allows the foundationalist to incorporate a valuable insight in coherentist theories of justification without having to subscribe to the stronger thesis that coherence confers justification.

4. Prominent Features of Mental State Defeaters

a. Newly Acquired State Defeaters and Newly Acquired Power Defeaters

Mental state defeaters may defeat beliefs at the time the defeater is acquired or they may do their defeating at some later time when they acquire the power to defeat. Bergmann (2006, pp. 155-57) designates the first a “newly acquired state defeater” and the latter a “newly acquired power defeater.”

Typically when we think of mental state defeaters we think of situations where a person *S* justifiably believes *p* at some time *t* but then at some later time *t** *S* acquires a mental state *d* (some new experience or belief) that causes *S*'s belief that *p* to be unjustified at *t**. Here *S*'s belief that *p* is unjustified from the time *S* acquires the mental state *d*. In the morning I hear the weather report and there's a prediction of showers late in the morning. Later in the morning I hear a pitter-patter against the window facing my backyard. Looking through my blinds, I see some dark clouds in the sky and water drops against my window. I justifiably believe at time *t* that it's raining outside. But suppose that several minutes later my wife walks in the front door (dry as a bone) and says that my next-door neighbor is spraying water over our fence on to the back of our house. It would seem that I'm no longer justified in believing that it's raining outside. At time *t* I was justified in this belief but at time *t** I'm no longer justified in this belief because I have acquired evidence at time *t** that defeats my prior belief. This is a newly acquired state defeater.

In other cases, though, a mental state *d* may be acquired at time *t* but not do its defeating work until some later time *t** when it acquires the *power* to defeat. Bergmann (2006, p. 156-57) designates this kind of defeater a newly acquired power defeater. Bergmann's illustration is helpful. My younger brother quietly tells me that when my sister comes into the room and informs everyone that my cousin Maggie is downstairs in the basement, this is really code for “Maggie is at her boyfriend's house.” As he explains, no one wants Maggie's father, who is present, to know that Maggie is at her boyfriend's house. My sister then enters the room and says she was just talking with Maggie downstairs, which I know really means that Maggie is at her boyfriend's house. As it happens, I already believe this because earlier in the day Maggie's boyfriend told me that Maggie would be visiting him at his house. So I have a justified belief that Maggie is at her boyfriend's house, even before my sister suggests this through code. Now suppose that shortly after the announcement, my older very reliable brother tells me that my younger brother was just trying to fool me with the code story. There was no plan for my sister to speak in code about Maggie. In this scenario, it looks like I acquire a mental state at a particular time that only subsequently acquires the power to defeat a belief of mine. I believe *B* (Maggie is at her boyfriend's house) at time *t* when I acquire the belief *M* (my sister has said that Maggie is in the basement), but the belief *M* does not defeat the belief *B* at time *t*. My belief *M* only gains the power to defeat my belief *B* after my older brother informed me that my younger brother was engaged in high jinx with me. This allows me to take my sister's comment as indicative of the actual whereabouts of Maggie, thereby defeating my prior belief Maggie is at her boyfriend's house.

Of course, in both the case of a newly acquired state defeater and a newly acquired power defeater the defeater may not be a complete defeater, that is, it may not render a belief wholly unjustified. While defeaters are normally thought of as rendering a belief unjustified or irrational, depending on the specifics of the evidential situation they might

merely render a belief less justified than it was before the acquisition of the defeater or before it acquired its defeating power. For example, suppose that when my wife tells me that our neighbor is spraying the backside of our house with his garden hose my wife has the kind of look she gets when she's trying to fool me about something. At the time, I can't fully accept what she says, but it's not obvious that she's trying to pull my leg. Perhaps her testimony in this case lowers the degree of justification for my belief that it's raining outside, rather than renders this belief wholly unjustified. So we should distinguish between complete and partial defeat/defeaters.

b. Diachronic Aspects of Mental State Defeaters

The above account of mental state defeaters construes them as mental states that defeat a belief at some particular time. This way of thinking about defeaters is naturally suggested by the correlated synchronic view of justification, namely of some person's being justified in believing p at some particular time t . But we can extend this view of defeaters by viewing their defeating power – like justification generally – through time or diachronically. (On the nature and significance of synchronic and diachronic justification, see Swinburne, 2001, pp. 152-91).

First, although mental state defeaters are naturally thought of as rendering unjustified (or less justified) a person's prior justified belief, mental states at some time t may also *prevent* a person from coming to hold a justified belief at some later time t^* . We might call this the forward-looking defeating potential of mental states. Suppose that my wife enters the house moments before I hear the pitter-patter and see the water drops against my window. She informs me about my neighbor's bizarre behavior of spraying the backside of our house. The subsequent perceptual evidence that would otherwise justify my belief that it's raining outside will not do so in this case. The potential justification conferring power of this evidence acquired at time t^* is antecedently neutralized by what I know or justifiably believe beforehand at time t . We might say that my wife's testimony constitutes a *preventative justification defeater*. More generally, at any given time t our experiences and set of justified beliefs will prevent us from being justified in holding some other belief(s) at some subsequent time t^* . Thus all mental states have some forward-looking defeating potential. Of course, we typically don't end up holding such beliefs (because we take them to be unjustified for us given the rest of what we believe), but if we did they would be unjustified by virtue of our other mental states.

Secondly, the defeating power of some mental state over an antecedently held belief can be said to continue into the future. Call these *continuing defeaters* (Bergmann, 2006, p. 158). The natural way of thinking about this is to take the case where someone continues to hold the defeated belief (or continues holding it with the same degree of firmness), despite the acquisition of a mental state defeater for the belief. Suppose that some of Kurtis' neighbors accuse Kurtis' wife Cathleen of having an affair with a married neighborhood man. Cathleen denies this and Kurtis justifiably believes that Cathleen is telling the truth. Later that day Kurtis sees Cathleen in a romantic embrace with a neighborhood man behind a tree in the local park. Kurtis has acquired a defeater for his belief that Cathleen is an honest wife, but through a variety of rationalizations he

continues to believe that Cathleen is an honest person. Kurtis' seeing Cathleen romantically involved with another man causes his belief in her honesty to be unjustified. Kurtis' memory of what he saw (or his belief that he saw it) continues to cause his belief in Cathleen's honesty to be unjustified, though he nonetheless persists with this belief. So here we have a case where a memory or belief state continues to make another belief – the subject persists in holding – unjustified. The defeater has continuing defeater power over a persisting belief.

Of course, the idea of a preventative justification defeater allows us to think of the defeating power of a mental state continuing into the future, even if the person gives up the defeated belief. Perhaps I give up my belief that it's raining outside after my wife tells me that my neighbor is spraying my house with a garden hose. In this case, at time t^* d (the awareness of my wife's testimony) is a defeater for a belief I had at the earlier time t but don't have any longer. Now at time t^* it makes no sense to speak of d as defeating my actual belief that it's raining outside, because I no longer hold this belief at t^* . But we can still speak of d 's continuing power to prevent me from forming the justified belief that it's raining outside.

c. Defeater-Defeaters

Mental state defeaters can of course be subsequently defeated by other mental states, and we can say that all mental state defeaters are continuing defeaters until they are defeated. That is, they continue to render a belief unjustified or less justified until their defeating force is neutralized. It's common to speak of mental states that defeat mental state defeaters as defeater-defeaters (Pollock 1986, pp.45-58; 1970; Plantinga, 1993a, pp. 231-37; 1993b, pp. 216-221; 1994; 1986). Suppose I justifiably believe T , Tom Grabit stole a library book. Now suppose I get a defeater D for the belief that T , namely Mrs. Grabit tells me that Tom is thousands of miles away and his identical kleptomaniac twin was at the library at the time in question. If I subsequently learn that Mrs. Grabit is a compulsive liar and deranged, then I have acquired a defeater D^* for the original defeater D . I have acquired a defeater-defeater. While D rendered my belief that T unjustified, D^* restores my justification for believing T .

Notice that in this particular example that D^* doesn't render my belief that D unjustified, even though it restores my justification for believing T . I'm still justified to believe D , namely that Mrs. Grabit said such and such. What is defeated here is the power of D to defeat my prior belief that Tom Grabit stole the library book. Take another example. Suppose I see what appear to be blue widgets coming down an assembly line. I believe that these are blue widgets. I then discover that the widgets are being illuminated with a blue light. This gives me a defeater for my belief that the widgets are blue. If I subsequently pick up a widget outside the range of the blue light, view it under normal lighting conditions, and see that it's blue, the defeating force of <these widgets are being illuminated with a blue light> is neutralized, but not in such a way that I cease to be justified or rational in believing that the widgets are being illuminated by a blue light. So when it comes to defeater-defeaters my justification for holding the originally defeated belief can be restored without causing the defeater against this belief to be an unjustified

belief. Defeater-defeaters might do that of course, but they need not. Perhaps I discover that what I thought was a blue light shining on the widgets is not a blue light at all or perhaps I learn that Mrs. Grabit actually did not say what I thought she said. In these cases the defeater-defeater causes my belief in the original defeater to be unjustified.

According to Plantinga (1986), some beliefs can, by virtue of their own degree of warrant, defeat defeaters that come their way. When a belief has this power, Plantinga designates it an intrinsic defeater-defeater against some ostensible defeater. I write a letter to the chair of my department trying to bribe him to write a highly exaggerated letter on my behalf for an NEH fellowship. The letter mysteriously disappears from the chairperson's office. I have a motive to steal it, the opportunity to do so, and I have been known to do such things in the past. Moreover, a reliable member of the department claims to have seen me hanging around the chairperson's office about the time the letter must have been stolen. Given the evidence, my colleagues believe that I stole the letter. Perhaps they are justified in believing this. However, *I* believe that I spent the day in the woods and so could not have stolen the letter. My memory belief has a great deal of nonpropositional warrant for me. So despite the counter-evidence, I'm justified to believe that I was in the woods and didn't steal the letter. Here it seems that the ostensible defeatee actually operates as a defeater-defeater. Plantinga of course isn't suggesting that an actually defeated belief restores warrant to itself by defeating an acquired defeater. It's not as if my belief that I didn't steal the letter was actually defeated at some point in time and its justification subsequently restored. The idea is rather that the original belief prevents or insulates itself from being defeated because the defeating *potential* of counterevidence is antecedently neutralized by the degree of warrant had by original belief. So I never actually acquire a defeater for my belief that I was in the woods or that the belief that I didn't steal the letter (Sudduth, 1999, pp. 180-82).

5. Variations on Mental State Defeaters

Advocates of mental state defeaters (and the corresponding no mental state defeater condition) differ on some crucial points regarding mental state defeaters.

a. The Epistemic Status of Defeating Beliefs

One of the issues of debate between adherents of [MSD] is whether beliefs that function as mental state defeaters must have some positive epistemic status to have defeating power, specifically if they are to defeat beliefs that do have some positive epistemic status. Plantinga (2000, pp. 364-65, 2002, pp. 272-75) contends that irrational and unwarranted beliefs can defeat beliefs that are (otherwise) rational and warranted. Suppose I believe that I'm made of flesh, blood, and bone. I then come to believe – due to some cognitive disorder - that my head is made of blown glass. According to Plantinga, given that I come to hold this second belief I now have a defeater for the prior belief, even if the defeater was formed by way of cognitive malfunction. In other cases, my belief may be rational but nonetheless unwarranted, and yet it might still function as a defeater for a warranted belief. Using another example from Plantinga (2000, pp. 363-65), suppose I believe that you were born in Yankton, South Dakota. Your uncle, whom

I believe to be a reliable person, told me this. My belief is warranted. But then one day you inform me in all seriousness that you were actually born in New Haven, Connecticut and you provide a reasonable explanation for why your uncle thinks otherwise. Absent any reason to suppose that you're trying to fool me or are delusional, I have a defeater for my belief that you were born in Yankton, South Dakota. However, suppose that your parents actually lied to you about where you were born. In that case, your belief that you were born in New Haven, Connecticut would not be warranted (given Plantinga's understanding of warrant), and neither would my belief that this is where you were born. So the defeater in this situation would be an unwarranted belief of mine. (Note that it also follows from Plantinga's account of defeaters that a belief *D* can defeat a belief *A* with no warrant, and that *D* can defeat a belief *A* that has more warrant than *D*).

Now in the above cases I acquire what Plantinga calls a "rationality defeater." By virtue of acquiring the defeating belief *D* I'm no longer rational in believing *A*. This is a consequence of an internal aspect of cognitive proper functioning, what Plantinga specifically designates *internal rationality*. Plantinga distinguishes between the proper functioning of our cognitive faculties "downstream" from experience (internal rationality) and the proper functioning of our cognitive faculties "upstream" from experience (external rationality) (Plantinga 2000, pp. 110-12). The former refers to the appropriate belief *response* to phenomenal imagery and doxastic experience, whereas the latter refers to proper functioning in the *production* of phenomenal imagery and doxastic experience. Internal rationality will include coherence among our beliefs and drawing the appropriate sort of inferences from what we believe. So to say that I have acquired a rationality defeater *D* for my belief *A* is to say that a certain doxastic response is called for *given that* I have a sensuous or doxastic experience of a certain sort. Perhaps I'm externally irrational in forming *D* (e.g., because I'm suffering from paranoia, dementia, or some kind of mental illness), but I'll still be internally irrational to continue holding *A* given that I hold *D*.

Alston (2002) has argued that Plantinga's position is counter-intuitive, and that only beliefs with positive epistemic status can defeat beliefs that have positive epistemic status, and a belief *D* can defeat belief *A* only if *D* has greater warrant than *A*. The efficacy of a defeater depends on the relative positive epistemic status of each of the beliefs being compared. Bergmann (2006, pp. 164-66) argues that Alston's rebuttal to Plantinga is plausible as an account of belief revision or how we ought to change our beliefs. Since Plantinga parses his own account of defeaters in this way, Alston's criticism is applicable to Plantinga's position. However, Bergmann maintains that Alston's argument doesn't undermine the notion that irrational or unjustified beliefs can defeat justification. My belief that I have hands is unjustified if I believe (however irrationally) that I'm a brain in a vat, even if it's more reasonable as a policy of belief revision to give up the belief that is less rational or less warranted.

b. Subjective and Objective Contours

Another issue, related to the first, concerns the relationship between *having* a mental state defeater and *believing* that one has such a defeater.

Plantinga (1994) suggests that, ordinarily at least, having a defeater involves one seeing or taking it that one's belief is defeated. But would this be sufficient for having a defeater?

Alston's criticism above entails that merely taking one's belief to be defeated isn't sufficient for defeat, because one might irrationally or unjustifiably take one's belief to be defeated. This is presumably the case when, due to my irrationally believing that my head is made of blown glass, I take my belief that my head is made of flesh, bone, and blood to be defeated. Alston and some other externalists would argue that only truth-conducively justified or reliably produced beliefs can be defeaters. However, since the truth-conducivity of grounds of belief and reliability of belief formation are not introspectively accessible facts, it is possible for an otherwise internalist no-defeater condition to be parsed with an external or objective component arising from the demand that defeaters be drawn from the subject's stock of justified beliefs or knowledge.

Internalists too may impose a similar requirement, so even if it's necessary that the subject take his belief to be defeated (in order to have an efficacious defeater), it will also be necessary that defeating beliefs have positive epistemic credentials of some sort. If my belief that Jack is a lifeguard is to be defeated by my belief that Jack can't swim, then the latter belief must be rational or justified. And for the internalist (unlike the externalist) *that* a belief has this kind of status will itself be a matter that is introspectively accessible. Moreover, the internalist will likely require that there be the appropriate kind of negative evidential relationship between the defeater and the defeatee. That is to say, if belief *d* actually defeats *S*'s belief that *p*, then *p* will at least not be likely given *d* and the relevant rest of *S*'s beliefs. *D* must sufficiently lower the evidential probability of *p*. If we suppose that criteria of inductive (and deductive) reasoning are introspectively accessible, then an internalist version of the no mental state defeater condition can be internalist in this additional respect. It can require the absence of a negative logical relation between *d* and *S*'s belief that *p*, where this is introspectively accessible and so can be determined upon reflection. (Swinburne, 2001, pp. 28-31).

Bergmann (2006, pp. 160-63), however, argues for a more subjective account of defeat, which he believes is at least suggested by both Plantinga and Pollock. On Bergmann's view, a person *S* has a defeater for his belief that *p* just if he consciously takes his belief that *p* to be defeated, and a person *S* takes his belief that *p* to be defeated just if *S* takes the belief that *p* to be epistemically inappropriate. For the latter, *S* must simply take himself to have good reasons for denying *p* or good reasons for doubting that the grounds of his belief that *p* are trustworthy, truth-indicative, or reliable. It isn't necessary that the person have what are *actually* good reasons for the negative epistemic evaluation of his beliefs. It is only necessary (and sufficient) that the person *take* himself to have such reasons, and Bergmann places no restriction on what kinds of considerations might play this role for the subject. So on Bergmann's view the no mental state defeater condition (as requirement for knowledge) is really a no *believed* defeater condition (Bergmann, 2006, p. 163). Bergmann's no defeater condition, then, is strongly internalist since one has introspective access to whether or not one takes a particular belief to be epistemically

inappropriate, even if there's no introspective access to either the justificational status of a defeating belief or the causal origin of one's taking a belief to be defeated.

c. Conscious and Reflective Mental State Defeaters

Since mental state defeaters include beliefs and beliefs may be occurrent or dispositional, it will be helpful to distinguish between conscious and reflective mental state defeaters (Bergmann, 1997a, pp. 116-121). There is a distinction between defeating experiences or beliefs of which one is aware at time t and defeating experiences and beliefs of which one is not aware at time t but of which one would become aware upon reflection. Similarly, there's a distinction between consciously taking one's belief to be defeated and this being something that one would do upon reflection. Accordingly, someone who advocates [MSD] may suppose that knowledge requires either the absence of a conscious defeater or the absence of a reflective defeater.

I said above (3.a) that some externalists advocate [MSD], specifically parsed in terms of the subject S not taking his belief that p to be defeated. Alston (1988b) appears to argue that the absence of a mental state defeater is not a necessary condition for knowledge. However, it's fairly clear that Alston has in mind a reflective defeater, not a conscious defeater, much less a person S 's consciously taking his belief that p to be defeated. Alston asks us to suppose that there is some person who has acquired substantial evidence that his sensory experience is a radically unreliable guide to his physical environment, that he's been the subject of a mad scientist's neurophysiological experiments for several years. So the subject justifiably believes that his senses are not to be trusted. However, as this person is about to cross a street he seems to see a truck heading towards him, and he forms the belief that a truck is approaching. His sensory perceptual system is working fine, and a truck is approaching. Alston says that in this scenario the person knows that a truck is approaching, despite having overriding reasons for supposing that his senses are not reliable. It would seem that the person has knowledge, despite having a mental state defeater. Crucial to Alston's account, though, is his claim that when the subject seems to see a truck approaching, he "momentarily forgets" his skepticism and acts accordingly. This makes it clear that the person in question does not consciously take his belief to be defeated when he sees the truck approaching. Rather, we have a reflective defeater, for the subject presumably would upon reflection take his belief to be defeated or epistemically inappropriate. So Alston's scenario can't plausibly be taken as a counter-example to a no conscious defeater requirement for knowledge, especially if this kind of requirement is parsed in terms of a subject not consciously taking her belief to be defeated.

The fact that a no conscious defeater requirement is widely subscribed to by both externalists and internalists counts in favor of its intuitive plausibility. But Bergmann (1997a, pp. 127-39) argues further that we have good reasons to reject the no reflective defeater requirement for knowledge. His argument is based on the premise that knowledge is incompatible with veritic epistemic luck but not evidential epistemic luck. Veritic luck refers to a person being lucky to believe what is true, given the evidence the person has. Evidential epistemic luck refers to a person being lucky to have the kind of

evidence she has. The Political Assassination, Unopened Letters, and original Tom Grabbit case discussed above (in 2.c) are arguably examples of evidential epistemic luck, whereas Goldman's Fake Barn case is an example of veritic epistemic luck. Bergmann argues that there are cases where a person has a reflective defeater for a belief, but the situation is analogous to cases of evidential epistemic luck. So we have reason for resisting the idea that knowledge requires the absence of a reflective defeater.

Here's Bergmann's example (Bergmann 1997a, p 136). Due to a strange cognitive disorder Chuck thinks that reports he hears between 4:15pm and 4:30pm are highly unreliable. On a particular day, Chuck's alarm clock wakes him up from an afternoon nap at 4:20pm. Immediately upon waking up Chuck hears noises outside his window. He looks and sees what appear to be city workers at work near a large hole in his front yard. One of the men tells Chuck that they are there to do work on the main waterline to Chuck's house, and that Chuck's wife was informed of this the day before. Chuck believes what he's told, and the man is telling the truth. However, if Chuck reflected on the matter, he would believe that the man's report was unreliable, for Chuck would have realized that he's being given this report between 4:15pm and 4:30pm and that reports he hears during this time period are unreliable. If Chuck reflected on the matter, he would consciously take it that his belief about what these men are doing is defeated. But Bergmann argues that most of us would be strongly inclined to say that in this scenario Chuck actually knows what the men in question are doing on his property, even though Chuck has a reflective defeater for this belief. Chuck is certainly lucky here not to have evidence against his belief, but in much the same way in some Gettier-type cases (e.g., Tom Grabbit case above in 2.c) the subject is lucky to have the evidence he does and not have other evidence (that is misleading) but it's not a matter of luck that the person believes what is true given the evidence he has.

6. Taxonomy of Defeaters and Formalities of Defeat

Having considered the distinction between propositional and mental state defeaters, something should be said about the formalities of such defeaters. It's fairly common for epistemologists to distinguish between two general ways beliefs may be defeated. There are defeaters that are reasons for supposing that p is false, and there are defeaters that are reasons that, if added to ostensible evidence for p , would sufficiently lower the likelihood that p is true. According to the first kind of defeater, we get reasons to believe the negation of p (or that p is false). According to the second, we simply lose our reasons for supposing that p is true. But let's look at the range of defeater-types.

a. Primary-Type Defeaters: Rebutting, Undercutting, and No Reason Defeaters

(i) A *rebutting defeater* for some belief that p is a reason (in the broad sense) for holding the negation of p or for holding some proposition, q , incompatible with p (Pollock, 1986, p. 38). Mary sees in the distance what appears to be a sheep in the field and forms the belief that there is a sheep in the field. The owner of the field then comes by and tells her that there are no sheep in the field. She has acquired what is commonly designated a rebutting defeater for her belief that there is a sheep in the field. She has acquired a

reason for supposing that there is no sheep in the field. Alternatively, she might have walked up to the object and discovered that it was actually a papier-mâché facsimile. Here she acquires a reason for believing something incompatible with her belief that there is a sheep in the field. These are of course examples of rebutting mental state defeaters. There can also be rebutting propositional defeaters. Perhaps Mary doesn't hear the owner of the field tell her that there are no sheep in the field, but he has mentioned this to several people in the neighborhood the day she believes there is a sheep in the field. There is a true proposition that counts against the truth of Mary's belief, even if it isn't a proposition she believes. (Of course, as noted above in connection with defeasibility analyses, there will be many true propositions that misleadingly count against the truth of beliefs).

(ii) An *undercutting defeater* for some belief that p is a reason (in the broad sense) for no longer believing p , not for believing the negation of p (Pollock, 1986, p. 39). More specifically, it is a reason for supposing that one's ground for believing p is not sufficiently indicative of the truth of the belief. A person enters a factory and sees an assembly line on which there are a number of widgets that appear red. Being appeared to red-widgetly, the person believes that there are red widgets on the assembly line. The shop superintendent then informs the person that the widgets are being irradiated by an intricate set of red lights, which allow the detection of hairline cracks otherwise invisible to the naked eye. Here the person loses his reason for supposing that the widgets are red, rather than acquires a reason for supposing that they are not red. Again, these are illustrations of undercutting mental state defeaters. There can also be propositional defeaters of the undercutting variety. The mere fact that the widgets are being irradiated with a red light would be one such example. Or suppose that Jason believes his tie is red. The fact that he is red-green colorblind might be a propositional defeater for this belief. The fact that someone is prone to perceptual hallucinations might be a propositional defeater for some range of sensory perceptual beliefs, and so forth.

(iii) A *no-reason defeater* is a reason for supposing that it's no longer reasonable to believe p given that (a) one has no reason for believing p and (b) the belief that p is the sort of belief that it's reasonable to hold only if one has evidence for p . (Bergmann, 1997a, pp. 102-103). For example, Johnny believes that if he dies he will immediately thereafter be turned into a zombie. Upon reflection he can't locate any reasons why he believes this, but he realizes that it's the sort of belief for which he ought to have some reason if he is rationally to believe it.

Now in each of these three cases (parsed in terms of mental state defeaters), the acquisition of a defeater makes it epistemically inappropriate to continue holding a particular belief B given that (i) there is evidence against B , (ii) reasons for B have become neutralized, or (iii) there is a recognition that one has no reasons at all for holding B though one ought to have such reasons. Consequently, a person's belief is no longer justified (or – in the case of partial defeaters – not as justified as it would be absent the defeater). If knowledge entails justification, each of these kinds of defeaters has the potential to defeat knowledge. If parsed in terms of propositional defeaters, then the

corresponding true propositions are such that they prevent an overall justified true belief from counting as knowledge.

b. Secondary-Type Defeaters: Defeaters for Grounds of Inferential Beliefs

There are also defeater-types that appear to be derived from (i), (ii), and (iii), and which apply specifically to cases where beliefs are based on other beliefs, that is, inferential or mediate beliefs.

(iv) A *rebutting reason-defeating defeater* is a rebutting defeater against a belief, c , where c is a ground or reason for the belief that p . Mark believes that his computer has a hardware problem that is causing several operation errors. He believes this because his wife tells him that Peter told her this and Mark knows that Peter is an expert on computers. Later, though, Mark discovers that it was not Peter but John who told his wife this, but Mark believes that John has little knowledge about computers.

Thinking of defeaters in terms of argument forms, Pollock (1986, pp. 38-39) distinguished between reasons that attack a conclusion (rebutters) and reasons that attack the connection between the premises and the conclusion (undercutters). Rebutting reason-defeating defeaters are distinct from both rebutting and undercutting defeaters in Pollock's sense. In the language of argumentation, they attack neither the conclusion nor the connection between the premises and the conclusion. A rebutting reason-defeating defeater is a species of rebutting defeater (as I defined it above), but it's a reason to believe the negation of a belief, c , that functions as the ground or reason of another belief p . In terms of argument forms, we can say that a reason-defeating defeater is a rebutting defeater against a premise in some argument. This kind of defeater is also distinct from Pollock's undercutting defeater. In the case of rebutting reason-defeating defeaters, it's not that the grounds fail to be indicative of the truth of Mark's belief that his computer has a hardware problem, but Mark comes to believe that one of his original grounds for holding this belief is false. Like undercutting defeaters, in acquiring a rebutting reason-defeating defeater we lose our reasons for supposing that the target belief that p is true. As a result, the grounds lose their power to confer justification on the target belief. However, this comes about by way of acquiring reasons for supposing that a ground of the target belief is false. (See Bergmann 1997a, pp. 99-103, for further discussion on the distinction between undercutters and reason-defeating defeaters).

(v) If we continue to think of defeaters and defeat in terms of argument structures then we can apply undercutting defeaters to more complex grounds for belief, where a belief that p is based on some further belief, q , that is in turn based on some other belief, r . An *undercutting reason-defeating defeater* for some belief that p is a reason for supposing that the grounds, r , for some belief that q fail to be sufficiently indicative of the truth of q , but where q is itself a ground for believing p . In terms of general logic, the premises of arguments are often themselves supported by reasons, thereby creating sub-arguments. Just as we can acquire reasons for the negation of a premise in an argument, we can acquire reasons for supposing that the premises of a sub-argument fail to be indicative of the truth of a premise in some main argument. As with rebutting reason-defeating

defeaters, we lose our reasons for believing the main conclusion, p , but here we do so by virtue of losing our reasons for believing a premise, q , rather than by acquiring a reason for denying the premise q .

(vi) A *no-reason reason-defeating defeater* is simply the application of the no-reason defeater to the grounds of an inferentially held belief. In (iii) a belief is defeated because it's not based on any reason but is the kind of belief that is reasonable only if there are reasons for it (or the person believes this is the case). However, even where some belief that p is based on the belief that q , the belief q may be such that *it* isn't based on any reasons but it would be unreasonable to hold the belief that q unless it's based on reasons.

7. Conclusion

I've outlined two general types of defeaters in this paper, propositional defeaters and mental state defeaters. The former are conditions (external to the perspective of cognizer) that prevent an overall justified true belief from counting as knowledge. The latter are conditions internal to the perspective of the cognizer (experiences, beliefs, withholdings) that cancel, reduce, or even prevent justification. Propositional defeaters are designed to address the problem of accidentally true belief, whereas mental state defeaters arise from the variable nature of justification. Inasmuch as justification is necessary for knowledge, mental state defeaters are capable of defeating knowledge. This leads to the viewpoint that knowledge requires the absence of any mental state defeater. So both kinds of defeaters ultimately relate to conditions of knowledge, and I've tried to develop each in connection with their larger epistemological territory.

I've examined the complexities that arise in developing no propositional defeater and no mental state defeater conditions for knowledge. The defeasibility theorist must select from among different criteria to locate the relevant range of true propositions that are genuinely indicative of a defect in justification that prevents knowledge. Advocates of mental state defeaters face a range of other issues, from choosing more or less subjective accounts of mental state defeaters, to choosing between conscious and reflective types of mental state defeaters for the no defeater condition for knowledge. I also considered synchronic and diachronic aspects of mental state defeat.

In the latter part of the paper I outlined a taxonomy of defeaters that highlights the difference between getting defeaters for beliefs and getting defeaters specifically for beliefs based on reasons of varying degrees of complexity. Here I pointed out several of the dynamics that emerge within the taxonomy of defeaters. One of the more important distinctions is between losing one's grounds for believing p and acquiring reasons for believing the denial of p (or for believing something incompatible with p). I also considered several ways in which a subject might lose his grounds for believing p . While some of these involve a subject becoming unjustified in holding to some reason, r , for his believing p , others amount simply to the subject's reasons, r , losing their power to confer justification on the target belief that p while the subject remains justified in believing r .

8. References and Further Reading

Alston, William. 2005. *Beyond Justification: Dimensions of Epistemic Evaluation*. Ithaca: Cornell University Press.

Alston provides a systematic analysis of various epistemic desiderata and their implications for revising our approach to the concept of epistemic justification.

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

Alston examines Plantinga's evolutionary argument against naturalism and offers criticisms of Plantinga's suggestion that an irrational belief can function as a defeater.

Alston, William. 1989. *Epistemic Justification*. Ithaca: Cornell University Press.

This is Alston's collection of previously published essays on substantive and meta-questions in epistemology, including essays on foundationalism and the concept of epistemic justification.

Alston, William. 1988a. "An Internalist Externalism." *Synthese* 74: 265-83. Reprinted in Alston, 1989, pp. 227-45. Page references are from reprint.

Alston develops a theory of epistemic justification that combines elements of externalism and internalism.

Alston, William. 1988b. "Justification and Knowledge." *Proceedings of the World Congress of Philosophy*, 5. Reprinted in Alston, 1989, pp. 172-82. Page references are from reprint.

Alston argues that justification (construed in both internalist and externalist ways) is not necessary for knowledge. The essay includes an argument for supposing that a person can know p even though she has a certain kind of mental state defeater for her belief.

Alston, William. 1986. "Internalism and Externalism in Epistemology." *Philosophical Topics*, 14: 179-221. Reprinted in Alston, 1989, pp. 185-226. Page references are from reprint.

Alston's examination of internalist and externalist approaches to justification.

Alston, William. 1983. "What's Wrong with Immediate Knowledge?" *Synthese*, 55:73-95. Reprinted in Alston, 1989, pp. 57-78. Page references are from reprint.

Alston critically examines various objections to “immediate knowledge” and argues that these objections rest on various implausible assumptions about the character of immediate knowledge.

Alston, William. 1976. “Has Foundationalism Been Refuted?” *Philosophical Studies*, 29: 287-305. Reprinted in Alston, 1989, pp. 39-56. Page references are from reprint.

Alston defends “minimal foundationalism” against the criticisms of foundationalism raised by Frederick L. Will and Keith Lehrer.

Annis, David. 1973. “Knowledge and Defeasibility.” *Philosophical Studies*, 24: 199-203.

Critical response to the defeasibility analysis provided by Lehrer and Paxson in Lehrer and Paxson, 1969, and which examines the nature of misleading or defective defeaters.

Audi, Robert. 1993. *The Structure of Justification*. Cambridge: Cambridge University Press.

Audi’s previously published essays on various topics in epistemology, including his development and defense of moderate foundationalism and the idea of “negative evidential dependence.”

Barker, John. 1976. “What You Don’t Know Won’t Hurt You.” *American Philosophical Quarterly*, 13: 303-308.

Barker attempts to tackle the Gettier problem in terms of a defeasibility analysis that distinguishes between genuine and misleading defeaters.

Beilby, James (ed). 2002. *Naturalism Defeated? Essays on Plantinga’s Evolutionary Argument against Naturalism*. Ithaca: Cornell University Press.

Essays discussing Alvin Plantinga’s evolutionary argument against naturalism, some of which discuss Plantinga’s notion of rationality defeaters.

Bergmann, Michael. 2006. *Justification without Awareness*. New York: Oxford University Press.

Bergmann defends an externalist theory of justification, which includes both a proper function and no mental state defeater requirement.

Bergmann, Michael. 2005. “Defeaters and Higher-Level Requirements.” *Philosophical Quarterly*, 55: 419-36.

Bergmann, Michael. 2000. “Deontology and Defeat.” *Philosophy and Phenomenological Research* 60: 87-102.

Bergmann argues that deontologism does not lend support to internalism. Essay provides several helpful observations on defeaters.

Bergmann, Michael. 1997a. "Internalism, Externalism, and Epistemic Defeat." (PhD Dissertation: University of Notre Dame).

Bergmann provides a detailed examination of the nature of defeaters and their relation to internalist and externalist theories of knowledge.

Bergmann, Michael. 1997b. "Internalism, Externalism, and the No-Defeater Condition." *Synthese*, 110: 399-417.

Bergmann argues that the no mental state defeater condition being necessary for warrant is compatible with externalist theories of warrant. Section 4 contains an analysis of externalists who endorse some version of the no mental state defeater condition.

Boonin, Leonard G. 1966. "Concerning the Defeasibility of Legal Rules." *Philosophy and Phenomenological Research*, 26: 371-78.

Boonin examines the meaning of defeasibility in law and its implications for legal analysis.

Chisholm, Roderick. 1989. *Theory of Knowledge*. 3rd edition. New Jersey: Prentice-Hall.

Chisholm provides an internalist response to the Gettier problem, as well as an account of defeasible justification influenced by defeasibility in moral philosophy. First edition: 1966.

Gettier, Edmund. 1963. "Is True Belief Knowledge?" *Analysis*, 23: 121-23.

Gettier's famous paper in which he argues that beliefs can be both true and justified and yet fail to constitute knowledge.

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

Goldman endorses a version of reliabilism with a no mental state defeater requirement for justification.

Goldman, Alvin. 1976. "Discrimination and Perceptual Knowledge." *Journal of Philosophy*, 73: 771-91.

Goldman discusses a causal theory of perceptual knowledge and defeasibility analyses of knowledge. The essay includes the famous “Fake Barn” scenario, a Gettier-type case initially suggested to Goldman by Carl Ginet.

Harman, Gilbert. 1973. *Thought*. Princeton, NJ: Princeton University Press.

Text contains Harman’s “Political Assassination” and “Unopened Letters” Gettier cases.

Hart, H.L.A. 1961. “The Ascription of Responsibility and Rights.” In Herbert Morris (ed), *Freedom and Responsibility: Readings in Philosophy and Law*. Stanford, CA: Stanford University Press, pp. 143-48. Originally published in *Proceedings of the Aristotelian Society*, 1948-49, 49: 171-94. Page references are from the reprint.

This is Hart’s classic discussion of the defeasibility of legal rules.

Klein, Peter. 1981. *Certainty: A Refutation of Skepticism*. Minnesota: University of Minnesota Press.

Klein presents his revised and detailed development of a defeasibility analysis of knowledge.

Klein, Peter. 1976. “Knowledge, Causality, and Defeasibility.” *The Journal of Philosophy*, 73: 792-812.

Klein, Peter. 1971. “A Proposed Definition of Propositional Knowledge.” *The Journal of Philosophy*, 68: 471-82.

Klein presents a defeasibility analysis of propositional knowledge to handle the intuition that knowledge cannot be accidentally true belief.

Lehrer, Keith and Paxson, Thomas. 1969. “Knowledge: Undefeated Justified True Belief.” *Journal of Philosophy*, 66: 225-37.

Influential early defeasibility analysis of knowledge in response to the Gettier problem, focusing on the problem of specifying the relevant sub-set of true propositions that are indicative of a defect in justification. The essay includes the widely discussed “Tom Grabbit” illustrations.

Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, MA: the Belknap Press.

An externalist account of knowledge that requires that the absence of a certain kind of mental state defeater, specifically that a person not believe that his belief does not track truth.

Plantinga, Alvin. 2002. "Reply to Beilby's Cohorts." In James Beilby (ed), 2002, pp. 204-75.

Plantinga responds to criticisms of his evolutionary argument against naturalism. His detailed comments on rationality defeaters are particularly relevant.

Plantinga, Alvin. 2000. *Warranted Christian Belief*. New York: Oxford University Press.

Plantinga applies his externalist theory of warrant and proper function to questions regarding the positive epistemic status of Christian belief. In chapter 11 Plantinga provides a more developed account of his view of rationality defeaters earlier introduced in Plantinga 1993a and 1994.

Plantinga, Alvin. 1996. "Respondeo" in Jonathan Kvanvig (ed), *Warrant in Contemporary Epistemology*. Lanham, MD: Rowman and Littlefield, pp. 307-78.

Plantinga responds to various criticisms of his externalist theory of warrant and proper function. Particularly relevant here is Plantinga's discussion of defeasibility analyses of knowledge in response to Klein and Swain, pp. 317-26.

Plantinga, Alvin. 1995. "Reliabilism, Analyses, and Defeaters." *Philosophy and Phenomenological Research*, 55: 427-64.

Plantinga, Alvin. 1994. "Naturalism Defeated." (Unpublished). Available on-line at <http://www.homestead.com/philofreligion/files/alspaper.htm>

An early version of Plantinga's evolutionary argument against naturalism in which he provides some detailed reflections on rationality defeaters, subsequently developed by Plantinga in Plantinga 2000.

Plantinga, Alvin. 1993a. *Warrant and Proper Function*. New York: Oxford University Press.

Plantinga's earlier discussion of rationality defeaters and the defeater system (pp. 40-42, 216-37) in the larger context of his theory of warrant as requiring the proper functioning of our cognitive faculties.

Plantinga, Alvin. 1993b. *Warrant: The Current Debate*. New York: Oxford University Press.

Plantinga articulates various inadequacies in contemporary internalist and externalist theories of warrant. The appendix examines Pollock's conception of defeaters.

Plantinga, Alvin. 1986. "The Foundations of Theism: A Reply." *Faith and Philosophy* 3, 3: 310-312.

Plantinga responds to Philip Quinn's criticisms of Plantinga's proper basicity thesis regarding theistic belief. Plantinga presents the idea of an intrinsic defeater-defeater.

Pollock, John. 1986. *Contemporary Theories of Knowledge*. Savage, MD: Rowman and Littlefield.

Pollock's account of justification utilizes a detailed account of mental state defeaters.

Pollock, John. 1984. "Reliability and Justified Belief." *Canadian Journal of Philosophy* 14, 103:114. Reprinted in Moser, Paul K. (ed). 1986. *Empirical Knowledge: Readings in Contemporary Epistemology*. Savage, MD: Rowman and Littlefield Publishers, pp.193-202.

Pollock discusses how the acquisition of reasons for supposing that a belief was unreliably produced defeat justification, but that this does not commit the epistemologist to a reliabilist theory of justification.

Pollock, John. 1974. *Knowledge and Justification*. Princeton, NJ: Princeton University Press.

Pollock, John. 1970. "The Structure of Epistemic Justification." *American Philosophical Quarterly*, monograph series 4: 62-78.

Article contains Pollock's early reference to two kinds of defeaters, Type I and Type II excluders, which later become rebutting and undercutting defeaters.

Shope, Robert. 1983. *The Analysis of Knowing: A Decade of Research*. Princeton, NJ: Princeton University Press.

Shopes provides an overview of a dozen or so early attempts to resolve the Gettier problem. Chapter two examines defeasibility analyses.

Steup, Matthias. 1996. *An Introduction to Contemporary Epistemology*. Upper Saddle River, NJ: Prentice Hall.

In chapter 1 Steup distinguishes between propositional defeaters (what he calls factual defeaters) and mental state defeaters (what he calls justificational defeaters) and considers their implications for various issues in epistemology.

Sudduth, Michael. 1999. "The Internalist Character and Evidentialist Implications of Plantingian Defeaters." *The International Journal for the Philosophy of Religion*, 45: 167-187.

Sudduth argues that Plantinga's notion of a "defeater system" (as a part of cognitive proper functioning) entails two significant evidentialist conditions for warranted belief in God.

Swain, Marshall. 1981. *Reasons and Knowledge*. Ithaca: Cornell University Press.

Swain attempts to address inadequacies in defeasibility analyses by combining a reliabilist indicator view of justification and a causal account of the basing relation.

Swain, Marshall. 1974. "Epistemic Defeasibility." *The American Philosophical Quarterly*, 11,1: 15-25.

Swain examines defeasible vs. indefeasible justification in relation to the Gettier problem and the analysis of knowledge.

Swinburne, Richard. 2001. *Epistemic Justification*. Oxford: Clarendon Press.

A development and defense of epistemic internalism, with chapters on Bayesian probability. Swinburne adopts a defeasibility analysis to handle the Gettier problem (pp. 192-200), but also incorporates mental state defeaters in his account of justification (pp. 28-31).

Kvanvig, Jonathan L. (ed). 1996. *Warrant in Contemporary Epistemology: Essays in Honor of Alvin Plantinga's Theory of Knowledge*. Lanham: Rowman and Littlefield.

A collection of essays on Alvin Plantinga's theory of warrant by prominent contemporary epistemologists. See especially articles by Peter Klein (pp. 97-130) and Marshall Swain (pp.131-146), both of whom address defeasibility analyses of knowledge in relation to Plantinga's theory of warrant.

Author Information:

Michael Sudduth

Email: msudduth@comcast.net

San Francisco State University