5

The Gettier Problem and the Value of Knowledge

The topic for this chapter arises out of two sources. In the previous two chapters, we discovered two properties of belief that are valuable beyond that of truth, the properties of subjective justification and intellectual virtuousity. Both of these properties are valuable by themselves, so that it is epistemically better for a belief to be subjectively justified than not and it is better for a belief to be virtuous than not. Moreover, these properties have a value not swamped by the presence of truth, so that it is better to have a subjectively justified true belief than one that is true but unjustified in this sense, and it is better to have a virtuous true belief than to have a true belief that is not virtuous. In the former case, what makes the property of subjective justification valuable in a way not swamped by the value of truth is that this property is a transparent mark of truth. So it has value because of its connection to truth, but its value is not swamped by the presence of truth, as are other properties such as reliability that also have value because of their connection to truth. Intellectually virtuous belief is valuable because when a belief has this property, the believer is due credit for having a true belief. If we understand the virtues in terms of truth-conduciveness,[1] then this property is valuable in part because of its connection to truth, but not in such a way that its total value is swamped by the value of truth.

Hence, the first source of motivation for the present chapter is the discovery of important epistemic properties other than truth. The other motivation is the Gettier problem, for neither of these properties is

1. For an argument against this construal of the virtues, see Jonathan L. Kvanvig, *The Intellectual Virtues and the Life of the Mind: On the Place of the Virtues in Contemporary Epistemology* (Totowa, NJ: Rowman & Littlefield, 1992), chapter 6.

sufficient, either individually or jointly, in the presence of true belief for knowledge. Thus, we must face the further Socratic query of what makes knowledge more valuable than true belief plus either or both of these properties.

I begin with a brief summary of the Gettier problem and some of the more important counterexamples in the literature. I then outline the basic approaches to the Gettier problem before turning to the question of whether any of these approaches offer the hope of explaining what makes knowledge better than subjectively justified true belief, virtuous true belief, and true belief that is both subjectively justified and virtuous.

THE GETTIER PROBLEM

The heart of the Gettier problem arises whenever the requirements other than truth for knowledge do not guarantee the presence of truth itself. I will call such a position "fallibilism." The position Gettier attacked is the position that knowledge is justified true belief, where justification can obtain even though it provides no guarantee of truth. In our case, the problem arises because neither subjective justification nor intellectual virtuosity is a guarantee of truth.

Gettier provided two counterexamples against the theory of knowledge he rejects.[2] The first example involved ten coins in the pocket of someone in the room, and the second involved a friend, Smith, being in Boston. In both cases, sufficient evidence to justify a belief is presented where the belief in question is false (either *Jones, who is in this room, has ten coins in his pocket* or *Smith is in Boston*). In each case, the person presented with the evidence reasons deductively from the false claim that is justified to a further claim (either *someone in this room has ten coins in his or her pocket* or *Smith is in Boston or Brown is in Barcelona*), where this further claim just happens to be true (either because somebody else in the room has ten coins in his or her pocket or because Brown just happens to be in Barcelona). Because the inference is a self-conscious and competent wielding of logical devices, any justification present for the premise ought to be transferred to the conclusion. If so, however, the conclusion is a justified true belief. The way in which it is true bears little resemblance to the path of discovery followed by the person in this

2. Edmund L. Gettier, "Is Justified True Belief Knowledge?" *Analysis*, 23 (1963): 121–3.

case, so it counts as a justified true belief where the truth of the belief is merely accidental, not related in an appropriate way to the justification that is present. Because it is accidental in this way, it does not count as knowledge, and hence knowledge cannot be identified with justified true belief.

These counterexamples attack a theory that is not in play at this point in our study, but it is not hard to adapt them to attack the theories that are in play. Consider first the claim that knowledge is subjectively justified true belief. Because subjective justification involves transparent marks of truth, and whether a mark of truth is transparent or not is such a subjective matter, there is nothing to prevent us from imagining the Gettier cases to involve an individual for whom all the justification presented and the inferences involved are transparent marks of truth. They may not be such marks for most people, but it is not impossible that they be such; and if they are, Gettier's examples adapted to such an individual count straightforwardly against the idea that knowledge is subjectively justified true belief.

Moreover, the same strategy shows that knowledge is not virtuous true belief. The individual in the Gettier cases could easily be a competent logician and an excellent handler of empirical evidence, thereby making his original belief and the inferences drawn displays of intellectual virtue. Even so, knowledge would not be present, so knowledge is more than intellectually virtuous true belief.

A final extension of the Gettier cases shows that knowledge is not intellectually virtuous and subjectively justified true belief. Simply let one individual have both the characteristics of the previous two paragraphs: She is both subjectively justified and intellectually virtuous in believing the concluding propositions of the Gettier cases. Still, she lacks knowledge, so knowledge is more than true belief plus the valuable properties elicited from our discussions in Chapters 3 and 4.

GETTIER-LIKE CASES

Here are four important test cases from the literature that any account of knowledge must address:

The Nogot–Havit Case: Nogot does not own a Ferrari, but he has provided Smith with overwhelming evidence that he, Nogot, owns a Ferrari. Smith comes to Nogot's office, where he talks with Nogot and the janitor, Havit. Because

Smith believes that Nogot owns a Ferrari and believes that Nogot is in the office, Smith infers that someone in the office owns a Ferrari. This belief is true because the janitor owns a Ferrari. Further, the belief is justified because it is knowingly inferred from other justified beliefs, the beliefs that Nogot owns a Ferrari and that Nogot is in the office. Yet, Smith does not know that someone in the office owns a Ferrari, even though he has a justified true belief that this is so.[3]

The Fake Barn Case: Henry is driving in the countryside with his son, teaching him the kinds of objects found in it. "That's a cow," says Henry; "that's a tractor," "that's a silo," "that's a barn," and so on. Henry has no doubt about the identity of these objects, and his vision is fully adequate and functional. In particular, Henry is aware of no grounds that would cast doubt on the last claim about a barn. Each of the identified objects has features characteristic of its type, Henry is reasonably careful in his identification, and there is little traffic to distract him. Unknown to Henry, the county he travels in is full of fake barns. These fakes look from the road exactly like barns, but are without back walls or interiors, quite incapable of being used as barns. The object Henry sees, however, is the lone real barn in the area, so Henry is correct in identifying it as a barn. Yet, if he had been looking at a fake, he would have mistaken it for a barn. So Henry does not know that it is a barn, even though he has a justified true belief that it is a barn.[4]

The Tim–Tom Case: Joe, the library detective, sees his good friend Tom take a book from the library and leave without checking it out. On the basis of perception and personal acquaintance, Joe justifiably believes that Tom stole a book. Joe informs the police officer of what he saw, but after talking to Joe the officer speaks to Tom's mother, who claims that Tom is out of town and that it may have been Tom's identical twin brother, Tim, who stole the book. Unknown to the officer, Joe, or anyone else, Tom's mother is lying. Tom stole the book, he was not out of town, and Tom does not have a twin brother. In spite of the fact that Joe has a justified true belief that Tom stole the book, the testimony of Tom's mother undermines his knowledge.[5]

The Assassination Case: A political leader is assassinated, but his associates, fearing a coup, decide to pretend that the bullet hit someone else. The false report appears on national television, stating that a secret service person has been killed in a failed assassination attempt. Before the announcement is made, however, a

3. Keith Lehrer, "Knowledge, Truth and Evidence," *Analysis*, 25, 5 (1965): 168–75.
4. Alvin Goldman, "Discrimination and Perceptual Knowledge," *Journal of Philosophy*, 73, 20 (1976): 771–91.
5. Keith Lehrer and Thomas Paxson, Jr., "Knowledge: Undefeated Justified True Belief," *Journal of Philosophy*, 66, 8 (1969): 225–37.

reporter files the correct story, which is reported in his newspaper. Jill buys a copy and reads the correct report, a report in a credible news source by a reliable reporter. Everyone else has seen both the newspaper article and the television report, and they are in a state of confusion. Because of the existence of evidence she does not possess that is possessed by everyone else, Jill does not know even though her belief is justified and true. How could she know because her belief is so dependent on her lack of generally available information?[6]

Some adaptation of these cases is required in order to make them threaten the account of the value of knowledge that we have achieved to this point. On that account, knowledge is valuable because it involves true belief, which is valuable, and involves one or both of the properties of intellectual virtuosity or subjective justification. The adaptations are not difficult, however; we need only include the relevant properties in each of the cases. Doing so creates no inconsistency in the cases and no greater inclination to think that knowledge is present once these properties are included.

The result of such cases is something we noticed earlier in our discussion of Swinburne's account of the value of knowledge. Swinburne defended the value of knowledge through an internalist conception of justification. His account of the value of that kind of justification is inadequate, as we saw, but we also noticed something equally important. Even if his account of the value of justification had been adequate, the preceding examples show that knowledge is not justified true belief and hence that the value of knowledge cannot be identified with the value of justified true belief. An adequate account of the value of knowledge must explain why it is more valuable than any subset of its constituents. If we assume that there is some property like justification that distinguishes knowledge from true belief, then an adequate explanation of the value of knowledge could be achieved by giving an adequate account of the value of justification. Because knowledge is more than justified true belief, such an explanation is only one part of a complete explanation. In addition, what is needed is an explanation of why knowledge is more valuable than justified true belief.

In the previous two chapters, we identified two properties slightly different from Swinburne's concept of internal justification that are valuable, but the lesson learned in discussing Swinburne's theory applies here as well. In order to understand fully the value of knowledge, we need an explanation of why knowledge is more valuable than subjectively justified

6. Gilbert Harman, *Thought* (Princeton, NJ: Princeton University Press, 1973), pp. 120–72.

true belief, more valuable than virtuous true belief, and more valuable than subjectively justified, virtuous true belief.

THE APPEAL TO ACCIDENTALITY AND A GENERAL CONCERN

Our success in finding valuable properties related to knowledge in the previous two chapters gives us hope that the same result may be forthcoming here, but I will argue that there is a general difficulty in attempting to explain the value of knowledge over the value of its components here that is new. I want to explain the general difficulty and then canvass particular attempts to provide a fourth condition for knowledge to show how these attempts fail to provide resources for addressing the general difficulty. The conclusion at which I aim is that the prospects are dim for an explanation of the value of knowledge arising out of the search for a solution to the Gettier problem.

I begin with the general problem. The heart of the Gettier problem involves some degree of slippage between what makes a belief true and what makes a belief justified (or intellectually virtuous), leaving the person in question with a justified belief that is only *accidentally* true. When epistemologists attempt to describe the general features of the problem they address, the concept of accidentality plays a central role along with other concepts such as "fortuitousness," "luck," and the like. This terminology leads to the guiding idea that knowledge is justified belief whose truth is no accident or the truth of which is neither lucky nor fortuitous.

These characterizations are on target, but they can be misunderstood. For they might be taken as a suggestion that we can arrive at an adequate account of knowledge by finding ways to rule out accidentality or luck or fortuitousness in having a true belief. It is instructive that epistemologists attempting to solve the Gettier problem have not made such attempts, for it is a mistake to think of the appeal to accidentality as supplying some theory that needs only minor tinkering in order to solve the Gettier problem. A more accurate picture of what epistemologists are doing when they cite accidentality as somehow central to the Gettier problem is that they are identifying some genus under which they believe they will find a species that provides a solution to the problem they seek to solve.

The idea that the appeal to accidentality is a theory only in need of refinement has not been pursued by epistemologists, and I think there is a very simple and compelling reason why it shouldn't be pursued. This idea

ignores the role a suitable, cooperative environment plays in the acquisition and possession of knowledge. We have found that an intellectually virtuous character is valuable on the basis of credit due to the believer for the accomplishment of true belief, and if we combine this approach with the idea of eliminating luck from true belief, we will begin to think of knowledge in terms of true belief achieved through the display of intellectual character excellent enough that no accidentality in the possession of true belief is present. Such an approach is hopeless, however, for explanations of true belief never appeal solely to the quality of our intellectual endowments. The environment itself must also be suitable for the operation of those powers. If we had infallible powers, things would be different; we would be in control of our epistemic destinies, and we could explain the difference between the value of knowledge and the value of true belief solely in terms of the value of our maximally excellent cognitive powers. But we do not have such powers, and hence the cooperation or suitability of the environment is required for a grasp of the truth to constitute knowledge.

Once we see that it is never solely in virtue of our cognitive powers that we find the truth, we must grant that a bit of fortuitousness is always present when we find the truth: Knowledge is always obtained at least in part by grace rather than totally by works.[7] So if it is fortuitousness that we hope to eliminate, we are hoping for something that cannot be had short of possessing infallible powers of discernment. I grant that such powers are surely desirable and valuable, but any attempt to account for the value of knowledge that appeals to the value of infallible powers of discernment will surely fail. For knowledge simply has little to do with infallibility.

These points do not by themselves give us a reason to think that the Gettier problem presents an insuperable obstacle to solving the *Meno* problem, for even if we cannot eliminate fortuity completely, there may still be kinds of fortuity that it is valuable to eliminate. Another point must be recognized as well, however. Given that knowledge is always obtained at least partially by grace, there will be elements of fortuity that we should not want to eliminate if knowledge is valuable. We cannot endorse the idea that any time any kind of fortuity is identified, it will be valuable to eliminate, for if we eliminate all fortuity, only infallible beings will have knowledge. Consider, for example, the fortuity of having been conceived,

7. I borrow this apt phrase from Robert J. Fogelin, *Pyrrhonian Reflections on Knowledge and Justification* (Oxford: Oxford University Press, 1994).

of having been born alive, of still being alive. There is also the fortuity each of us experiences of having many more beliefs about a particular locale rather than other locales (because of where we are physically located in space-time). Moreover, fortunate accidents bless us constantly, from the failure to acquire diseases that degrade our cognitive equipment to the failure of our enemies to carry out the work of the evil demon by breaking connections between what is rational to believe from our own perspective and what is true or likely to be true.

We might say, then, that the best approach to the Gettier problem would be one that identifies a kind of fortuity that, as it were, wears its disvalue on its sleeve. That is, it would be most useful for solving the *Meno* problem if the kind of accidentality eliminated by the fourth condition for knowledge were one that is intuitively disvaluable. Such a proposal would mimic the proposal that justification is required for knowledge in addition to true belief. Justification is, by its normative or evaluative nature, a valuable property. It takes no transcendental deduction or complex theoretical framework to establish this point. As such, the property of justification appears, prima facie at least, to be well suited for use in an explanation of the value of knowledge over that of true belief.

We have seen that the explanation of the value of justified true belief over that of mere true belief requires complexities not accounted for by this prima facie perspective. Yet the point remains that it is the intuitive value of justification that undergirds the search for an adequate and useful solution to the difficulties encountered. If it were not obvious from the outset that justification is a valuable property, the problems encountered would have provided sufficient reason in themselves to look elsewhere to account for the value of knowledge.

The same points should drive discussion of the Gettier problem and the prospects for solving the particular version of the *Meno* problem that arises only after noticing that knowledge cannot be identified with justified true belief. If we can identify some property that is intuitively valuable to eliminate, then there is hope for a solution to the problem of the value of knowledge. If, however, the best we can do is to offer a proposal that is so gerrymandered and ad hoc that we cannot identify this solution with any recognizably disvaluable kind of luck, accidentality, or fortuitousness, except to say that it is just that kind eliminated by whatever closes the gap between justified true belief and knowledge, we face a serious problem. For if such a characterization is all that could be provided, we would have no reason whatsoever for thinking that a solution could be found to the problem of the *Meno* on the basis of the value of knowledge's constituents.

This point is a straightforward implication of the twin desiderata on a theory of knowledge I have been stressing. On the assumption we have been making in the past several chapters (that the value of knowledge is in some way a function of the value of its parts), the need to account for both the nature and value of knowledge requires that we identify a fourth condition that not only yields a counterexample-free account of knowledge but also provides some basis for explaining the value of knowledge over the value of its constituents. The point I have been arguing is that this dual requirement can only be met when the identified fourth condition is something more than a gerrymandered, ad hoc way of avoiding counterexamples to one's account of knowledge. The identified fourth condition must identify some species of accidentality or luck, where such an identification provides a basis for defending the value of eliminating such luck.

The labyrinthine complexity of solutions to the Gettier problem is well known, and it is common to characterize the search as one where the complexity involved is ad hoc and gerrymandered in excess. Williamson, for example, argues that the very complexity of analyses of knowledge is incompatible with the value knowledge possesses:

> Even if some sufficiently complex analysis never succumbed to counterexamples, that would not entail the identity of the analyzing concept with the concept *knows*. Indeed, the equation of the concepts might well lead to more puzzlement rather than less. For knowing matters; the difference between knowing and not knowing is very important to us. Even unsophisticated curiosity is a desire to *know*. This importance would be hard to understand if the concept *knows* were the more or less ad hoc sprawl that analyses have had to become; why should we care so much about *that*?[8]

I do not wish to endorse all that Williamson says here, but only to point out the inverse proportionality he sees between the complexity and ad hoc character of an account of knowledge and the usefulness of such an account in an explanation of the value of knowledge. His remarks present a nice summary of the point I am arguing. We might put the point in terms of being pulled in two different ways when addressing the Gettier problem. On the one hand, the variety of cases in which one can fail to know and yet have a justified true belief inclines one toward more complex, ad hoc, and gerrymandered proposals. On the other hand, the felt need

8. Timothy Williamson, *Knowledge and Its Limits* (Oxford: Oxford University Press, 2000), pp. 30–1.

to address the question of the value of knowledge over its subparts leads one toward simpler proposals in which the value of the condition is intuitively obvious. The twin desiderata on a theory of knowledge, the desiderata of accounting for both the nature and value of knowledge, threaten in this way to become the Scylla and Charybdis that sink the project entirely. I want to argue that this threat is real in the following sense: The Gettier problem shows that no component-based account of the value of knowledge will be successful. This conclusion leaves open the possibility that the value of knowledge can be accounted for in other ways, but it closes the door on the idea that we can account for the value of knowledge in terms of the value of its subparts.

APPROACHES TO THE GETTIER PROBLEM

I will canvass some of the more popular attempts to solve the Gettier problem to see how they fare with regard to this tension between the desiderata of the nature and value of knowledge. As might be expected, most such theories founder by focusing on the complexities needed to account for the nature of knowledge, thereby undermining any ability to explain knowledge's value. I want to begin, however, with a theory having the opposite problem, a simplistic version of the relevant alternatives approach.[9] I will argue that this version begins with a property that is intuitively valuable but has no hope of success in accounting for the nature of knowledge because it refuses to define the concept of "relevance."

On a relevant alternatives approach, the difference between knowledge and justified true beliefs is determined by whether one would be immune to error in alternatives to the actual situation. For example, in Goldman's fake barn example, an alternative to the actual situation is one where one is referring to a fake barn rather than a real barn when saying, "That's a barn."

This theory handles the fake barn case quite well, but it also risks implying global skepticism if we consider the alternative situation to the actual one in which Descartes's evil demon is operative. In order to avoid this skeptical consequence, this approach introduces the qualifier "relevant" and holds that the evil demon scenario is not a relevant alternative to the actual situation. The pressing issue for this approach is to specify

9. For an early version of the view, see Alvin Goldman, "Knowledge and Perceptual Discrimination," *Journal of Philosophy*, 73 (1976): 771–91.

what makes a situation relevant, and here relevant alternatives theorists have had little of help to say. The most simplistic version of the view would simply rely on our intuitive understanding of the concept of relevance, claiming that no more precise theoretical specification is needed.

Such a theory is well suited to addressing the issue of the value of knowledge. Immunity from error is itself a good thing, and it would be hard to argue that one should prefer such immunity in *irrelevant* alternatives to immunity in *relevant* alternatives. Whether this value could withstand the scrutiny needed to provide a complete answer to the question of the value of knowledge remains to be seen, but the theory provides some hope of such. It provides such hope by identifying a property with obvious evaluative dimensions and in this way follows the strategy of addressing the question of the value of knowledge by identifying evaluative features of knowledge not present in mere true belief.

Those schooled in the esoterica of the literature on the Gettier problem will be impatient with this defense of such a simplistic version of the relevant alternatives theory. For without some clarification of the concept of relevance, this approach is a nonstarter for addressing the problem of the nature of knowledge. It is important to recognize explicitly the significance of the intuitive concept of relevance, however. For the evaluative nature of this concept is precisely what one would wish for if one were focusing primarily on the question of the value of knowledge. It is unfortunate that the simplistic version of this approach has no similar hope of adequately addressing the question of the nature of knowledge. As such, it provides a good example of how a myopic focus on one of the two desiderata on a theory of knowledge leaves one with a theory holding little promise for success.

I do not wish to be misunderstood here as claiming that all versions of the relevant alternatives approach are of precisely the same sort as this simplistic version, for there are ways in which relevant alternatives theorists have gone beyond a simple appeal to the concept of relevance. One of the more interesting attempts is to appeal to contextualism here, claiming that what is relevant varies by context. This answer can be successful in addressing the question of the value of knowledge only if what is valuable covaries with this concept of relevance, and we will have a chance to examine this position in detail in a later chapter.

A relevant alternatives theorist might resort to a circular account of knowledge, claiming that a relevant alternative is one in which the failure to arrive at the truth is a failure that undermines knowledge. Such an

approach is an obvious nonstarter, but it is interesting to compare this strategy with others that we will see later. This approach offers a substantive start on the problem of the value of knowledge and founders by offering an obviously circular nonstarter on the issue of the nature of knowledge. Other approaches to the Gettier problem end up doing the opposite, I will argue. These approaches offer a substantive start on the question of the nature of knowledge, but can offer only a circular nonstarter on the issue of the value of knowledge. They identify a condition on knowledge that can only be seen as valuable by assuming that the value of knowledge must be dependent on the value of its constituents. Such, we will see, is the lot of theories that focus myopically on the question of the nature of knowledge.

We will need to put this point on hold for the time being, however, for some ground clearing is necessary before we can take a look at the most promising approaches to the Gettier problem, approaches that share with the relevant alternatives theory this feature of offering false hopes. The ground clearing is necessary because of the interplay between the twin desiderata on a theory of knowledge, of accounting for both the nature and value of knowledge. Many of the approaches to the Gettier problem suffer from such deep defects in the attempt to account for the nature of knowledge that it is pointless to ask what use they have in addressing the problem of the value of knowledge. The goal of this ground clearing will be to isolate those approaches that have some prospect of success in providing an account of knowledge that is, at a minimum, free from counterexample in order to assess whether such approaches can be of any use in accounting for the value of knowledge. I'll begin with some general classificatory remarks about the kinds of approaches available for addressing the Gettier problem.

Some of the earliest attempts to explain the difference between knowledge and justified true belief emphasize the role that falsehoods play in yielding cases of justified true belief that are not cases of knowledge. Some claim, for example, that one cannot have reasoned through any false steps in arriving at the belief in question or that one's reasoning cannot involve these false steps essentially.[10] This approach emphasizes actual beliefs in the reasoning process, whereas other theorists employ the notion of "falsehoods" without requiring actual beliefs. Roderick Chisholm, for example, has suggested that knowledge requires evidence that confirms no

10. See, for example, Harman, *Thought*, chapter 9.

falsehood,[11] and Ernest Sosa claims that knowledge cannot epistemically presuppose any falsehoods.[12]

Defeasibility theorists provide a different approach. According to them, knowledge requires that justification be undefeated, where a defeater is a claim that, if believed, would undermine whatever justification is present. For defeasibility theorists, this necessary condition for being a defeater is not itself sufficient, and one difficult task for such a theory is to determine which subclass among those claims that satisfy this necessary condition is defeaters. One prominent suggestion is that a defeater is also a proposition that the individual in question is justified in believing to be false.[13]

Other approaches emphasize concepts such as "causality," "reliability," or "conclusive reasons," where the last idea is clarified in terms of counterfactuals. In the first group is Mark Steiner, who claims that knowledge of *p* requires that the sentence that expresses *p* be used in a causal explanation of the person's belief that *p* is true.[14] Reliability theories must appeal to reliability in a way different from reliability theories of justification, on pain of having to identify knowledge with justified true belief and thereby affirm the presence of knowledge in the cases presented earlier. One example of a different, and stronger, appeal to reliability to distance knowledge from justified true belief is offered by Adam Morton, who holds that knowledge has to be produced or sustained by a process that produces or sustains no false beliefs, and that is such that there is an explanation in terms of the laws of nature and the facts of the case, of why on this occasion a true belief results.[15]

Conclusive reasons approaches are clarified in terms of counterfactuals. Fred Dretske's original conclusive reasons account required that if the proposition were false, one would not believe it, and if one did not believe the proposition, it would not be true.[16] L. S. Carrier suggests the following: a person *S* knows that *p* only if (a) the reasons for *S*'s believing *p* are such that in *S*'s circumstances, if it were not the case that *p* then *S*

11. Roderick Chisholm, *Theory of Knowledge*, 2nd edition (Englewood Cliffs, NJ: Prentice-Hall, 1977), chapter 6.
12. Ernest Sosa, "Epistemic Presupposition," *Justification and Knowledge: New Studies in Epistemology*, George Pappas, ed. (Dordrecht: Reidel, 1979), pp. 79–92.
13. See, for example, Lehrer and Paxson, "Knowledge."
14. Mark Steiner, "Platonism and the Causal Theory of Knowledge," *Journal of Philosophy*, 70 (1973): 60.
15. Adam Morton, *A Guide Through the Theory of Knowledge* (Encino and Belmont, CA: Dickenson, 1977), p. 58.
16. Fred I. Dretske, "Conclusive Reasons," *Australasian Journal of Philosophy*, 49 (1971): 1–22.

would not believe *p*, and (b) the reasons for *S*'s believing *p* are such that, in *S*'s circumstances, if *S* were not of the belief that *p*, it would not be the case that *p*.[17] In his comprehensive summary of the search for a fourth condition for knowledge, Robert Shope defines the conclusive reasons approach as one involving either Carrier's condition (a) or one of the following:

> (a′) There is some subset *H* of existing circumstances that are logically [and causally] independent of the truth of *p*, such that unless *p* were the case, *S* would not believe *p*; or (a″) . . . unless *p* were the case, *S* would not have the reasons *S* does for believing *p*.[18]

Such conclusive reasons theories are close cousins, at least, of the relevant alternatives theory. If we adopt something like the standard semantics for counterfactuals, then a conclusive reasons approach requires that we be immune from error in certain close counterfactual situations. It is easy to use this language to explain the relevant alternatives theory. What it requires is the same immunity from error in situations that are relevantly close to the actual situation.

Each of these approaches begins from an intuitively attractive starting point, responding as they do to certain subsets of counterexamples that have arisen in response to Gettier's original counterexamples. My goal is to show how the promise each approach provides disappears as refinements to the approach are introduced. In some cases, the promise of the approach founders on unanswerable difficulties regarding the nature of knowledge, and I want to move past these approaches first in order to focus on those approaches that appear to have some possibility of correctly explaining the nature of knowledge.

Consider first the approaches that focus on the role falsehoods play, approaches that claim that the absence of knowledge is a result of reasoning that contains falsehoods, or evidence that contains or presupposes or confirms falsehoods. Such approaches suggest that knowledge might be important because it insulates us from error, beyond the object of belief itself, and such extended insulation from error is certainly a valuable property to have. For if truth itself is valuable, then more truth will be valuable as well.

17. L. S. Carrier, "An Analysis of Empirical Knowledge," *Southern Journal of Philosophy*, 9 (1971): 6.
18. Robert K. Shope, *The Analysis of Knowing: A Decade of Research* (Princeton, NJ: Princeton University Press, 1983), p. 122.

The problem, however, is that knowledge doesn't require such insulation from error. The first of these approaches to be abandoned is the view that knowledge requires reasoning through no false steps. Such a claim is of no help with examples such as the fake barn case, in which no reasoning takes place at all. So the only approaches with any promise at all are those that emphasize bodies of evidence that contain, presuppose, or confirm falsehoods. These approaches, though they have a bit more promise, cannot withstand scrutiny either. First, among the lessons of the preface and fallibility paradoxes are that we nearly always have evidence in support of some false propositions. The lottery paradox begins by imagining a fair lottery with a thousand tickets in it. Each ticket is so unlikely to win that we are justified in believing that it will lose. So we can infer that no ticket will win. Yet we know that some ticket will win. In the preface paradox, authors are (sometimes) justified in believing everything in their books. Some preface their book by claiming that, given human frailty, they are sure that errors remain, errors for which they take complete responsibility. But then it appears that they justifiably believe both that everything in the book is true and that something in it is false.

There are a number of attempts to solve these paradoxes, but no such attempt can deny the possibility of knowing that some ticket will unexpectedly win the lottery and knowing that undetected errors remain in our best work on a topic. If such knowledge is possible, however, the evidential base for such knowledge will support some falsehoods. To know that errors remain undetected in a work requires knowing something of our fallible natures and knowing the contents of the work. And such knowledge must also involve seeing each of the claims in the book as true, for that is what it is for the errors to remain undetected. Similar remarks apply to the lottery case: To know that some unknown ticket will win requires knowing something about the nature of the lottery, but it also involves strong evidence for each ticket that it will lose, for that is what is involved in the case of the lottery in knowing that some ticket will unexpectedly win the lottery. So knowing that some ticket will unexpectedly win requires epistemic support for some falsehoods, and knowing that undetected errors remain in a work also requires epistemic support for falsehoods. So even though these types of insulation from error look like nice properties to have, they are not properties implied by knowledge.

There is another lesson that the preface paradox teaches, for a natural extension of that paradox raises the fallibility paradox. Let the book in question be a compilation of all of our beliefs. The preface statement

then becomes an assertion of our own fallibility, that among our beliefs are undetected falsehoods. For those of us aware of our own fallibility, we have knowledge that presupposes some falsehoods. So once again, the paradoxes teach that knowledge does not require some blanket insulation from error that would be exceedingly nice to possess.

Second, the possibility of statistical knowledge undermines such an approach as well. A statistical sample can deviate in statistically significant ways from what it confirms about a population and still be used to gather knowledge about that population. For example, a sample might give us knowledge that most swans are white even though the actual percentage of white swans is significantly different from what our sample confirms. In such a case, the statistical sample grounds some knowledge but also confirms some falsehoods. Suppose our sample has 80 percent white swans, and, being somewhat sophisticated statistically, we know that the range for two standard deviations is 76 to 84 percent, that is, there is a 95 percent chance that the actual population is in that range, given the randomness of our sample. I use this information to conclude that most swans are white and acquire knowledge on this basis. The sample confirms other claims – for example, that the actual population is between 78 and 82 percent white, because the likelihood of this claim is fairly high given my sample (not 95 percent, but, say, 90 percent). Suppose, then, that the population is actually 83 percent white. Then I know that most swans are white on the basis of evidence that confirms a falsehood for me.

If the actual population were quite skewed from my sample (suppose it is only 60 percent white), then we would need to grant, I think, that the sample can't give inductive knowledge. Furthermore, in the preceding case, I assume that if a probability for a claim is known to be .9, that's enough for confirmation. One might choose to set the threshold for justification higher, but the same conclusion can still be employed that knowledge can be present when one's evidence confirms some falsehoods.

The lesson is that approaches that emphasize the role that falsehoods play in Gettier problems will have to be restricted in order to avoid such difficulties for this approach. Once defenders of this approach back away from the claim that knowledge issues a blanket protection against falsehoods related in some special way to one's body of evidence, they will need to appeal to some feature of this set of falsehoods *different from their falsity* to explain why immunization from these particular falsehoods is valuable. No such approach of this type has attempted such an explanation, and I know of no promising approach along

these lines (in part because every approach along these lines is subject to counterexample[19]).

A second approach to the Gettier problem emphasizes concepts such as causality, reliability, or conclusive reasons characterized in terms of counterfactual claims. We have already seen reasons for thinking that reliability approaches succumb to the swamping problem, so approaches that employ such concepts will be unsuccessful here as well in explaining the value of knowledge. Moreover, standard causal theories of knowledge simply are not general enough to be adequate to all the kinds of knowledge that we have, including mathematical knowledge and moral knowledge. Steiner's approach cited earlier, which tries to salvage the generality of the causal approach, requires that sentences play a causal role in belief formation in order for knowledge to be present (in particular, he requires that the sentence that expresses *p* be used in a causal explanation of why the belief that *p* is true), but that approach simply cannot work. Some animals and prelinguistic children have knowledge, but sentences play no causal role in the explanation of the truth of their beliefs.

Perhaps what Steiner intends is not that the sentence plays a causal role, but rather that we, the explainers, must use that sentence in providing an explanation. Still, such a theory implies the existence of explainers, possibly distinct from those who hold the beliefs in question. Some worlds contain both explainers and believers, and some do not; so this attempt to salvage full generality for the causal approach cannot work either. Nor will it work to make the appeal to explainers conditional – that if a successful explanation were carried out, it would appeal to a sentence that expresses *p*. Just think of the attempt to explain the possibility of God knowing that he has created a world in which no linguistic items exist. The only worlds in which an explanation could be carried would be worlds in which the claim in question would be false, and hence the explanation would be unsuccessful.

Among this group, we are left, then, only with the conclusive reasons approach, which emphasizes what would happen, epistemically, in close counterfactual situations. This approach in terms of counterfactuals and the defeasibility approach provide the most promising approaches to the Gettier problem. The defeasibility theory is formulated in terms of what happens to justification when additional evidence is taken into account, for a defeater is a claim that, in conjunction with the evidence that justifies

19. See ibid., for a catalog of such attempts and counterexamples to each.

the belief, fails to justify that belief. In the remainder of this chapter, I want to focus on these more promising approaches to the Gettier problem. I want to argue that neither approach has much hope of accounting for both the nature and value of knowledge.

COUNTERFACTUAL AND DEFEASIBILITY APPROACHES TO THE GETTIER PROBLEM

Consider first how the defeasibility approach fails regarding the value of knowledge. At first glance, this approach offers much promise regarding the question of the value of knowledge, for it is a valuable to have an epistemic standing that is not defeated by any additional information. The difficulty is that this explanation of what is valuable is too strong, for knowledge doesn't require immunity from defeat by just *any* additional information, for one of the lessons of the literature is that not all defeaters undermine knowledge.[20]

How is this possible? That is, how can there be additional information that defeats justification but doesn't undermine knowledge? Consider a minor variant on the earlier Tim–Tom case. In that case, what defeats Joe's justification for believing that Tom stole the book is the testimony of Tom's mother. In particular, it is her claim that Tom was out of town and that Tom has a twin brother, Tim, that defeats Joe's justification. But now imagine the testimony also including an admission of lying. That is, suppose that Tom's mother, just after claiming that Tom was out of town and that he has a twin brother, says, "Oh, my gosh, what am I saying? I'm tired of trying to protect him. I recant this poor excuse of an alibi for Tom." Then the testimony of Tom's mother does not undermine Joe's knowledge, even though it contains a defeater of the justification that Joe has.

So even if we value undefeated justification, that gives us no reason to value knowledge over justified true belief, for knowledge does not eliminate defeaters. Part of what makes the concept of defeat significant is what it implies about further learning. If your justification is undefeated, then you could pursue further investigation in an unlimited fashion and never encounter information that justifies changing your mind. You would have the assurance that further investigation would continue to confirm exactly what you currently believe, giving one clear license

20. To my knowledge, Peter Klein first saw this. His own theory of the distinction is presented in *Certainty* (Minneapolis: University of Minnesota Press, 1981).

to stop the investigation if one chooses. Such assurance is lost, however, once we note that not all defeaters undermine knowledge.

The language used to describe the two kinds of defeaters I have been discussing is noteworthy and instructive. In cases where a defeater is present that does not undermine knowledge, the defeater is said to be a "misleading" one. Knowledge can then be understood in terms of justified true belief that is subject to no nonmisleading, or "genuine," defeaters. This language signals, I believe, an implicit recognition of the need to address the question of the value of knowledge, for these terms are clearly value-laden.

The difficulty for the defeasibility approach is precisely the one we encountered with the relevant alternatives approach earlier. The concept of relevance, too, serves to indicate the need to address the question of the value of knowledge, but fails to give an adequate account of the nature of knowledge unless it is explained in terms other than to say that an alternative is relevant if what happens in it can undermine knowledge. In the same way, if one's explanation of the distinction between a genuine and a misleading defeater were that the first undermines knowledge and the second doesn't, we would again have no successful addressing of the question of the nature of knowledge. In both cases, resources for explaining the value of knowledge are retained at the cost of circularity in the account of the nature of knowledge.

So if the defeasibility theory is an advance over the relevant alternatives theory, it can only be so by offering some more substantive account of the distinction between misleading and genuine defeaters. To their credit, defeasibility theorists have devoted considerable attention to this issue. The problem, I will argue, is that their attempts do not track any distinction in value, as the language of genuine and misleading suggests.

For example, one might try to redeem the defeasibility approach by claiming that when you are subject only to defeaters that do not undermine knowledge, you could learn everything epistemically relevant to your belief and still be justified in believing what you currently believe. This reply is at the heart of defeasibility approaches that appeal to the power of overriders to explain why some defeaters are misleading and others are not. When a defeater is misleading, there is a further piece of information that, together with one's original justification and the misleading defeater, still provides justification for one's belief. Moreover, this hierarchy of defeat and overriding is not limited to two levels but can go indefinitely high, so that one's understanding of knowledge is in terms of justifications for which any defeaters are ultimately overridden.

To see why this approach does not go very far in addressing the question of the value of knowledge, compare this hierarchy of defeaters and overriders with cases in which one's justification is simply defeated. In this latter case, it is still true that one's belief is correct and that one would still be justified in believing what one believes were one to learn everything epistemically relevant to what one believes. It is for this reason that one can define omniscience in terms of being justified in believing p if and only if p is true.[21] So we have two cases, one of knowledge and one not, but in both cases, there is information that, if learned, would undermine our justification. Furthermore, there is additional information beyond this defeating information that, if learned, would restore our justification. In the one case, these additional pieces of information are so related that knowledge is present. In the other case, they are not. Yet, nothing in the account gives any reason to suggest that additional pieces of information so related that knowledge is present constitute something of special value, over and above additional pieces of information so related that knowledge is not present. Hence, the defeasibility approach based on the hierarchy of defeaters and overriders is unhelpful in explaining the value of knowledge in terms of the value of its constituents.

Peter Klein has a slightly different proposal, one relying on the distinction he formulates in terms of "effective" and "initiating" defeaters. An effective defeater is one that breaks the chain of confirmation for one's belief. An initiating defeater is one that renders plausible some effective defeater, and such a defeater is misleading when the chain between it and the effective defeater depends essentially on misinformation independent of the evidence a person has. As Klein puts it, "The defeating effect of the initiating defeater is essentially parasitic upon misinformation which does not depend upon any false proposition in E_s [the set of evidence for the person in question]."[22]

There are concerns about this account that arise from the point of view of both the nature of knowledge and the value of knowledge. First, it is critical for the success of Klein's account that misleading defeaters, by themselves, do not undermine chains of confirmation. For example, in the Tim–Tom case, the testimony of the mother is supposed to be misleading evidence when it is part of a pattern of protective behavior of which the police are fully aware and regarding which there is a long

21. See Jonathan L. Kvanvig, *The Possibility of an All-Knowing God* (London: Macmillan, 1986), for a full defense of this account of omniscience.
22. Klein, *Certainty*, p. 148.

history. For Klein's account to succeed, the claim that the mother so testifies to the police cannot itself be an effective defeater. If this account is to succeed, that information cannot, by itself, undermine the confirming power of perceptual evidence and background acquaintance with the person who stole the book. The difficulty, from the point of view of the nature of knowledge, is that this restriction is not obviously satisfied. It is true that the testimony of the mother confirms other propositions that also defeat the confirming power of one's evidence regarding who stole the book, propositions such as *Tom has a twin brother whom I cannot distinguish from Tom, and this twin was at the library yesterday.* Klein's position requires, however, that only such propositions as the latter undermine the confirming power of one's evidence. I think that view is implausible. The contrary testimony of people deemed to be in a position of greater authority on a particular subject is, by itself, sufficient to undermine the confirming power of one's evidence. A sufficient explanation for abandoning a complex claim about probability is that my friend, Daren, an expert on probability theory, says the claim is false. That is all the reason I need for abandoning the claim, and it is a reason for abandoning the claim because in normal circumstances it is sufficient, in itself, to undermine the confirming power of whatever evidence I had for the claim.

My claim here may suggest a strongly nonholistic conception of confirmation, but it does not require it. Whether one thinks of the confirmation relation as a two-place relation between propositions or as a three-place relation between two claims and some background information, the point still holds. Expert testimony, by itself, is enough to undermine confirmation (given normal kinds of background information).

This objection to Klein's proposal is a concern over its adequacy as an account of the nature of knowledge. Concerns over its adequacy to help account for the value of knowledge are equally pressing. There are several ways, on his account, that defeaters can be present and yet not undermine knowledge. Consider again Klein's characterization of misleading defeaters quoted earlier:

> The defeating effect of the initiating defeater is essentially parasitic upon misinformation which does not depend upon any false proposition in E_s.

In attempting to account for the value of knowledge, we need to consider some contrasting kinds of defeaters. One kind is a "one-step-removed" defeater: information that renders plausible some other information that undermines one's evidence. We can contrast such a defeater with a "direct" defeater, a bit of information that simply undermines the confirming

power of one's evidence. If we use Klein's account to explain the value of knowledge over that of justified true belief, we might try to maintain that it is better to be subject to a one-step-removed defeater than to a direct defeater (because, on his account, one-step-removed defeaters are less likely to undermine knowledge than are direct defeaters). Such a proposal is akin to saying that second cousins are better than first cousins. From some perspectives, that may be true: Second cousins are perhaps less likely to ruin one's reputation by association. From other perspectives, it is simply not obvious. My selfish genes, it is said, find more at stake in the welfare of first cousins than second cousins, for example. Similar points can be made regarding Klein's distinction in types of unpossessed information.

There is, of course, a circular way to get the desired explanation of the value of knowledge: We could identify one kind of unpossessed information as more valuable because of its association with knowledge and then cite that value as the basis for explaining the value of knowledge over that of its subparts. If we are careful to avoid such obvious circularity, however, no account will be forthcoming, for there is no reason to think that immunity from direct defeaters is preferable to immunity from defeaters once removed.

One might try to explain the value difference here in terms of the possibility of confirming falsehoods, that a once-removed defeater may confirm falsehoods and a direct defeater won't, but that proposal will not work. Almost any (contingent) falsehood can confirm a falsehood, as we saw earlier in this chapter. The only relevant difference here is whether the confirmation of falsehoods is the kind that allows or prevents the certification of a knowledge claim, but we can't use that information to infer that one kind is better than the other without endorsing a blatantly circular account of the value of knowledge.

There is one more feature of Klein's proposal that bears noting. The proposal that "The defeating effect of the initiating defeater is essentially parasitic upon misinformation which does not depend upon any false proposition in Es" distinguishes between misinformation that depends on false claims in one's body of evidence and misinformation that does not depend on false claims in that evidence. The first kind is "false-evidence-dependent," the second not. When a defeater is parasitic on false-evidence-dependent misinformation, it cannot undermine knowledge; when a defeater is not parasitic on misinformation that is false-evidence-dependent, it can.

It is hard to see this distinction as anything more than gerrymandering needed to prevent counterexamples to one's account of the nature of

knowledge, and it is easy to side with Williamson in remarking, "Why should we care about *that*?" The distinction between these kinds of defeaters tracks no intuitive difference in value, leaving us with an account of the nature of knowledge incapable of helping to explain the value of knowledge. As noted before, there is a circular account available, one that presumes on the value of knowledge to explain the significance of the proposed distinction, but we have rehearsed the error of this way enough already.

Klein's proposal is instructive regarding the pattern to be found among defeasibility theorists and more generally among the more sophisticated approaches to the Gettier problem. The initial formulation shows signs of sensitivity to the need for giving an explanation of the value of knowledge, for the initial formulation normally cites a property of belief that would be valuable to have. After the initial formulation, however, emendations are introduced with no apparent eye at all to the issue of the value of knowledge. The procedure appears to presuppose that the issue of the value of knowledge has already been solved by giving an initial formulation sensitive to that issue, so that emendations can be made to the initial formulation that ignore the question of the value of knowledge. Of course, no one reasons explicitly in this way, but the pattern of refinement for the initial approaches has all the flaws of theorizing based on this reasoning.

A more likely explanation concerns the myopia that sets in when one is trying to solve an exceedingly difficult problem, and Klein's language is instructive in this regard. Just prior to offering his characterization of which defeaters fail to undermine knowledge, Klein says, "But *there is a common feature* in all of the cases in which the initiating defeater is misleading."[23] This language betrays a focus solely on the issue of the nature of knowledge, for the point of looking for a common feature among a group of cases is to find a criterion for dividing them that is immune from counterexample. It is no wonder, once such myopia takes over, that the result is a condition on knowledge with nothing to commend it for the task of explaining the value of knowledge.

It may be thought that such myopia is present from the start when dealing with the Gettier problem, but I remain unconvinced. The point is not all that important, but it is instructive to notice that the usual approaches to the Gettier problem are presented in value-laden terminology: relevant

23. Ibid; italics mine, added to highlight the instructive phrase.

alternatives, undefeated justification, conclusive reasons, and so on. Even if the theorizing pays no explicit attention to the question of the value of knowledge, this language shows the significance of the question and some (perhaps only implicit) interest in addressing it. The difficulty raised by the Gettier problem soon comes to dominate, and proposals become qualified with no eye whatsoever to the problem of the value of knowledge.

Once one becomes sensitized to the problem of the value of knowledge and its particular incarnation regarding the Gettier problem, it is easy to spot flaws in other versions of the defeasibility theory. For example, consider that version proffered by John Pollock and Richard Swinburne, aimed at accounting for the assassination case presented earlier, in which information generally available undermines any claim to knowledge. Swinburne says:

> I concur with Pollock that the difference lies in what 'we are socially expected to be aware of'. 'We are expected to know what is announced on television' . . . (I have replaced Pollock's phrase 'expected to know', question-begging in this context, by 'expected to believe'.)[24]

Swinburne notes that we cannot give a noncircular account of knowledge if we use the phrase "expected to know" and so prefers to formulate the position in terms of what we are expected to be aware of or to believe.

One problem with this approach is that it does not generalize well to other cases involving the difference between misleading and genuine defeaters, such as the difference between the two Tim–Tom cases discussed earlier. In one of these cases we have only the mother's testimony, and in the other we have her testimony plus an admission that she is lying and wishes to retract her testimony. Whatever the explanation of the difference between these cases, it has nothing to do with what we are socially expected to be aware of or to believe.

Furthermore, once we are sensitive to the need to account for the value of knowledge, we will question the usefulness of this approach, even if we had thought the approach was counterexample-free. We will wonder why it is important or valuable to possess information we are expected to be aware of, questioning how our contemporaries could be so unfailingly right about which information is important to possess and which isn't.

24. Richard Swinburne, *Epistemic Justification* (Oxford: Oxford University Press, 2001), pp. 191–2. The reference to Pollock's view is from *Contemporary Theories of Knowledge* (Totowa, NJ: Rowman & Littlefield, 1986), p. 192.

Immediately after noting his agreement with Pollock, Swinburne expresses some worries about the approach:

> But, of course, such expectations vary with the society, and give rise to large border areas with respect to which experts may well differ about what is socially expected or about which they agree that expectations are diverse. There may well be societies in which everyone consults a suspect's mother before reaching judgments about guilt. And how 'nationwide' would the television coverage need to be, in order to undermine Jill's claim to knowledge?[25]

Swinburne uses these considerations about the variability of expectations to introduce a discussion of contextualism, which may rescue this approach. These considerations suggest the need for a strong form of relativism about the concept of undefeated justification, a form of relativism whose weaknesses are not mitigated by an appeal to contextualism. One of the most difficult problems for any version of relativism regarding evaluative concepts is how infallibility could plausibly be vested in the attitudes of peoples or groups. If relativism is offered as an account of what is morally good and bad, the question is how a society could be incapable of error about such matters. If relativism is offered as an account of when certain defeaters undermine knowledge and when they do not, a similar question is pressing. Once we have focused on the question of the value of knowledge and are looking for an account of the fourth condition for knowledge that helps explain why knowledge is more valuable than mere justified true belief, we will want an answer to the question of how people's expectations could infallibly track this difference in value. There is simply no reason whatsoever to think that expectations can't be off-base about what is valuable in this respect, just as there is no reason to think that groups of people can't be mistaken about what is morally right and wrong.

The appeal to contextualism at this point is a nonstarter. I will discuss the contribution contextualism might make to the question of the value of knowledge in later chapters, but the point to note here is that if contextualism gives us nothing beyond the relativizing of questions of value to the attitudes of one's peers or any other group, it will be no help in addressing the problem of the value of knowledge. As we will see, contextualism comes in much more plausible forms, leaving room for contextualists to incur no obligation to solve this knotty problem for relativism. Contextualists are free to agree that from the perspective

25. Ibid., p. 197.

of the question of the value of knowledge, an account of the difference between misleading and genuine defeaters in terms of expectations of peers tracks no important difference between kinds of unpossessed information.

So, to turn away from specific kinds of defeasibility theories to the general approach that such theories take, we encounter once again a bald and unqualified approach that suggests a possible explanation of the value of knowledge over subsets of its constituents. Yet, the bald and unqualified approach yields an inadequate account of the nature of knowledge, so qualifications are necessary. The urged qualifications eliminate the attractiveness of the approach for explaining the value of knowledge, however, even if they raise the prospects for such an approach to succeed as a counterexample-free explanation of the nature of knowledge. The sad reality is that the hopes for satisfying the dual requirements on a theory of knowledge once again seem reciprocally related: As hope for satisfying one requirement rises, hope for satisfying the other diminishes, leaving us with an epistemology inadequate for explaining either the nature of knowledge or its value.

The last remaining approach is the conclusive reasons approach or, more generally, approaches that attempt to elucidate the nature of knowledge in terms of certain counterfactuals relating the claim believed with the holding of the belief. Labeling the approach as a conclusive reasons approach is a little misleading, because versions of this approach do not appeal to the concept of a reason at all but instead use counterfactuals containing only the concepts of belief and the truth and falsity of such, counterfactuals such as *If the claim in question were false, one would not believe it* and *In similar situations in which one believes the claim, it is true.*

To see why the language of reasons is out of place in this approach, note that which possible situations are closest to the actual situation is often an objective feature outside our understanding or grasp of a situation, and thus outside of, independent of, the reasons we possess. Consider a case where one has strong evidence for a claim, where one has checked and rechecked some simple arithmetical sums on a sheet of paper, confirming that all the work on that sheet is correct. In some such cases, the counterfactuals in question will be true, and in other cases with precisely the same quality of reasons, the counterfactuals will be false. For the reasons themselves do not *make* the counterfactuals true or false. A more natural suggestion is that the counterfactuals are true or false, depending on the nature of the persons involved. That nature might be the kind of nature that would result in the acquisition of the reasons in question, so reasons

do play a role, perhaps, in some way, but it is not the quality of reasons but rather the quality or character or nature of the person having the reasons that carries explanatory weight. That is, the reference to quality of reason is not nearly as important as, say, the reference to the person's ability to track the truth of what is believed.[26] The counterfactuals in question express the power or ability of a person with respect to the truth, because the person is such that the claim in question would not be believed if it were false, nor would one fail to believe the truth in close counterfactual circumstances, slightly different from the actual circumstances, in which the claim is true. With Ernest Sosa, we can call these two conditions "sensitivity" and "safety."[27] A belief is sensitive when it is abandoned in conditions where its content is false, and a belief is safe when it is held only when it is true.

Unfortunately, this approach is not as successful as one would like. First, note that the truth of the counterfactuals in question is of little interest. A benevolent demon might devote himself to making the appropriate counterfactuals true so that some person's beliefs are technically safe and sensitive; another, malevolent demon might make those same counterfactuals false regardless of the quality of my evidence or the content of one's intellectual character. Hence, the truth or falsity of these counterfactuals alone is of no special interest to us, so if their truth is what makes the difference between justifed true belief and knowledge, then knowledge is of no special importance either.

There is a reply here, that it is not the truth of the counterfactuals that interests us, but the *explanation* of their truth. We want the counterfactuals to be true in virtue of capacities or abilities we have in finding the truth, so even though the mere truth of the counterfactuals may not be important or valuable, their truth in virtue of our powers is. If the counterfactuals are true because of the activity of some demon, they don't interest us much; but if they are true in virtue of our cognitive powers, they interest us, and properly so.

The question we must ask, however, is what all this has to do with knowledge. It is not hard to see that knowledge is not found simply by adding this condition to those already determined to be valuable. First, there are counterexamples to the claim that sensitivity is a requirement

26. Robert Nozick, *Philosophical Explanations* (Cambridge, MA: Harvard University Press, 1981).

27. Ernest Sosa, "Plantinga on Epistemic Internalism," *Warrant in Contemporary Epistemology*, Jonathan L. Kvanvig, ed., (Totowa, NJ: Rowman & Littlefield, 1996), pp. 73–86.

for knowledge. Consider, for example, Goldman's dog example.[28] I look at a dachsund and know that it is a dog. If the animal confronting me weren't a dog, it would be a wolf (by whatever details of nature one wishes to include to make this particular claim true). I am not very good at distinguishing wolves from dogs, we might suppose, so if I were not currently confronted with a dog, I would still believe that what I was confronted with is a dog. Still, my failure to be able to distinguish wolves from dogs should have no bearing on the present case, for I am not at all confused about whether a dachsund is a dog.

For another example, consider a case of knowledge by induction. June drops a trash bag down a garbage chute at her apartment building. A short time later, the bag is in the basement garbage room. June knows about the structure of the garbage chute and has often gone to the basement to retrieve something from her garbage that she hadn't meant to throw away. On the basis of past experience and relevant background knowledge, June knows on this occasion that her garbage is in the basement. But her belief is not sensitive. In the extremely unlikely event that her trash bag does not end up in the basement, it would be because it became snagged on the wall of the chute, and if it did, June would still believe that her garbage is in the basement. Nonetheless, June now knows, on inductive grounds, that her garbage is in the basement.[29]

Sosa's safety clause, which requires that June would believe that her garbage is in the basement only if her garbage would be in the basement, fares no better with this case. First, note that we cannot read this safety clause in accord with the standard semantics for counterfactuals where any subjunctive is true when it has a true antecedent and a true consequent. Doing so would result in judging every true belief to be safe. And if every true belief is safe, then it is obvious that knowledge is not true belief that is both justified and safe – Gettier's own examples would undermine the view.

What Sosa has in mind regarding the truth conditions of this safety requirement is what goes on in close possible worlds. In particular, we are to consider close worlds in which the antecedent is true and see if the consequent is true in those worlds as well. For Sosa's account to succeed, he will have to maintain that no worlds are close enough to affect the assessment of the safety condition if such worlds contain an extremely unlikely event (such as the garbage not making it to the bottom of the

28. Goldman, "Discrimination and Perceptual Knowledge," p. 779.

29. Ernest Sosa, "Skepticism and Contextualism," *Philosophical Issues*, 10 (2000): 1–18.

chute). Put slightly differently, the truth of the safety condition is not affected by counterfactual situations in which extremely unlikely events occur. For if close worlds can contain unlikely events, then one close world would be one where June believes her garbage is in the basement when in fact it is hung up somewhere in the chute.

Such a requirement, the requirement that close worlds cannot contain extremely unlikely events, forces unacceptable conclusions in lottery situations. It forces one to hold that one can know that one's ticket will lose. The belief that one's ticket will lose can be justified if the lottery is large enough, but we should balk at the idea that such a belief can count as knowledge. If someone says that he knows his ticket will lose, he should be reprimanded: He is justified in being very confident that it will lose, but he does not know that it will lose.

Such a belief is nonetheless safe on the interpretation of safety required to render the safety requirement immune from the garbage chute example. On that interpretation, close worlds cannot contain extremely unlikely events. If so, however, then the belief that one's lottery ticket will lose is safe. It is true that a person would believe that his or her ticket lost only if it did in fact lose, if the interpretation of this conditional excludes worlds in which extremely unlikely possibilities, such as one's ticket winning the lottery, are excluded.

One might suggest, as does Greco, that we interpret safety a bit more weakly, so that a belief is safe when usually it would be held only if it is true.[30] Such a weakening makes the account even more susceptible to counterexamples, however. Consider a 1,000-ticket lottery and construct increasingly large disjunctions claiming that certain tickets win as follows: One of tickets 1–500 will win, one of 1–501 will win, one of 1–502 will win, and so on. At some point in this sequence, before reaching the claim that one of 1–1,000 will win, we will reach the threshold at which safety, on this interpretation, will have been reached, implying that it is possible to know that such a claim is true. We must reject such a claim, however, for we can construct an incompatible set of propositions of the same likelihood, all having the same claim to knowledge. Suppose the threshold is achieved at the 75 percent level, so that the belief that one of tickets 1–750 will win is a safe belief. Then so is the belief that one of tickets 2–751 will win, and so is the belief that one of tickets 3–752 will

30. John Greco, "Knowledge as Credit for True Belief," *Intellectual Virtue: Perspectives from Ethics and Epistemology*, Michael DePaul and Linda Zabzebski, eds. (Oxford: Oxford University Press, forthcoming), p. 4 typescript.

win. If we take the set of all combinations of 75 percent of the tickets and consider beliefs that each member of this set will win, we get an inconsistent set of beliefs that are nonetheless all safe by this weakened account of safety. Not all members of this set can be true, and yet every member of the set is in precisely the same epistemic condition on the interpretation of safety under consideration. The proper conclusion to draw is that this epistemic condition is insufficient for knowledge.

Timothy Williamson also has argued against a sensitivity requirement,[31] endorsing instead something like a safety requirement, claiming that when one knows, one cannot easily be mistaken – the very language used by Sosa to introduce the concept of safety. He gives a counterexample to sensitivity that contains language that provides the basis for learning a more general lesson. He says:

> I tend slightly to underestimate the distances I see. When I see a distance of twenty-one metres I judge it to be less than twenty metres, although when I see a distance of twenty-three metres I do not judge it to be less than twenty metres. This may mean that when I see a distance of nineteen metres and correctly judge it to be less than twenty metres, I do not know it to be less than twenty metres. It surely does not mean that when I see a distance of one metre and correctly judge it to be less than twenty metres, I do not know it to be less than twenty metres. . . . Suppose that a mark on the side of a ship is one metre above the waterline. . . . I judge by sight whether the mark is less than twenty metres above the waterline. Let *p* be the proposition that the mark is less than twenty metres above the waterline. If *p* had been false, I might still believe *p*. I believe *p* insensitively. Surely I can still know *p*, because I believe *p* on quite different evidence from that on which I would have believed it had it been false.[32]

This example is important for both sensitivity and safety theories, for Williamson's example is easily adapted to form a counterexample to safety as well as sensitivity. Williamson would still have knowledge in the case he describes even if his estimates of distance were much more unreliable. Suppose, for instance, that he judged anything over five meters to be no less than twenty meters. He can still know that the mark that is in fact one meter above the waterline is less than twenty meters above the waterline. His belief is not sensitive, but neither is it safe: In most cases where he judges the mark to be less than twenty meters above the waterline, he is wrong.

31. Timothy Williamson, *Knowledge and Its Limits* (Oxford: Oxford University Press, 2000), chapter 7.
32. Ibid., pp. 159–60.

Williamson's discussion contains an explanation of why such examples pose problems: The evidence would be different from the evidence on which the actual belief is based. So if any of the approaches that employ counterfactuals is going to succeed at identifying the nature of knowledge, the antecedent of such counterfactuals will have to include a clause requiring sameness of evidence. Failure to do so will leave such theories open to Williamson-like counterexamples.

Other refinements are possible as well. Nozick, for example, wants to restrict the antecedent of the sensitivity requirement so that the same method is employed.[33] As before, however, refinement occurs at the expense of attention to the question of the value of knowledge. The question we must always ask is whether the refinement proposed makes a difference in the realm of value. For example, suppose that a justified true belief is not safe or sensitive when these conditions are defined with no restriction on the method employed. Suppose further that a justified true belief is both safe and sensitive when these conditions are defined with a restriction on the method employed (or, if one prefers, with a restriction to ensure sameness of evidence). Terms like 'safe' and 'sensitive' are value-laden terms and seem like the kinds of properties it would be good for a belief to have. These terms, however, are merely placeholders for certain counterfactual conditionals, one set being restricted by the method employed and the other set not. As far as I can see, the only way to arrive at the conclusion that it is more valuable to satisfy one such set of counterfactuals is by first identifying that set as the one connected with knowledge and assuming that knowledge is valuable. If so, however, we have precisely the same difficulty here as I highlighted with the relevant alternatives approach. We have, that is, no resources for explaining the value of knowledge in terms of the value of its constituents, because the identified constituent has no claim to value except in virtue of some supposed connection with knowledge, which is assumed to be valuable.

So we arrive at the same point with the more promising approaches to the Gettier problem that we began with concerning the relevant alternatives approach. In each case, such approaches offer something of value that might be used to explain the value of knowledge, but each such approach faces immediate difficulty concerning the nature of knowledge. Counterexamples to the initial formulation of the approach force alterations in the approach, and the alterations are guided exclusively by concern

33. Nozick, *Philosophical Explanations*, pp. 179ff.

over the nature of knowledge, resulting in emendations of the original suggestion that appear entirely ad hoc from the point of view focusing on the question of the value of knowledge. That is, if we devote attention not only to the adequacy of such approaches in terms of the nature of knowledge but also to the adequacy of such approaches in terms of the value of knowledge, no such emendations have much to recommend them. As the prospects rise for providing a counterexample-free account of the nature of knowledge, the prospects sink for providing an account of knowledge in terms of the value of its constituents.

CONCLUSION

Two conclusions are warranted by our discussion. Extant approaches to the Gettier problem offer no basis for explaining the value of knowledge over and above the value of true belief, subjective justification, and the display of virtuous intellects. The failure of these approaches gives a strong inductive argument for thinking that success is unlikely.

Sometimes such inductive evidence is compatible with optimism about ultimate success, as may be the case regarding scientific theories. Even though we have a scientific track record of failed theories, that record may also contain signs of progress so that the inductive argument for thinking that present and future theories will all be false fails. Here, though, the evidence points otherwise, leaving little ground for optimism here. When we look carefully at the variety of failed attempts to satisfy the twin desiderata concerning the nature and value of knowledge, we do not find signs of progress. We find, instead, a repeated pattern in which progress with respect to one desideratum is balanced by greater weakness with respect to the other. The heart of the Gettier problem concerns the presence of a certain kind of accidentality or luck, and our discussion provides a long and sustained basis for thinking that any value found in eliminating this kind of luck can only be found by assuming first the value of knowledge.

We have, therefore, a strong argument against the approach to the question of the value of knowledge pursued in the past four chapters: the approach that attempts to explain the value of knowledge in terms of the value of its constituents. The difficulties encountered in attempting to explain the value of knowledge through the value of its constituents show that we need a different approach. If knowledge is valuable, it is so on the basis of factors other than the value of its constituents. I turn to such alternative attempts in the next two chapters.