REPRESENTATIVE CONCEPTS: HOW TO ANALYZE KNOWLEDGE AS TRUE BELIEF IN THE FACE OF GETTIER COUNTEREXAMPLES

Modie Christon Smith

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Philosophy.

Chapel Hill
2006

Approved by
Advisor: William G. Lycan
Reader: Dorit Bar-On
Reader: Ram Neta

ABSTRACT

MODIE CHRISTON SMITH: Representative Concepts: How to Analyze Knowledge as True Belief in the Face of Gettier Counterexamples
(Under the direction of William G. Lycan)

Gettier counterexamples purport to show that justified true belief is insufficient for knowledge and, *ipso facto*, that true belief is insufficient for knowledge. I develop a strategy that the proponent of the true belief analysis of knowledge can deploy to explain away Gettier counterexamples, i.e., to show how the subject in them can merely appear to believe something truly without knowing it. I suggest that the proposition that appears to be truly, non-knowingly believed in a Gettier counterexample is actually not believed at all. Rather, the subject believes a closely related, false proposition. The subject cannot believe the true proposition, I propose, because of the special nature of certain of his or her concepts—what I call "representative concepts." The reason the subject appears to believe the true proposition is that we use a true sentence to ascribe the belief in the false proposition to him or her.

CONTENTS

Introduction

There are two broad classes of counterexamples to the thesis that knowledge is merely true belief (henceforth, "TB").[1] The first class is more traditional and comprises what we may call "justification counterexamples." A justification counterexample is one in which a person appears to truly believe something without knowing it, and the reason she appears not to know it is that she lacks good evidence for it or good reasons to believe it—in short, the belief is not justified. The elements of the second, more recent class we may call "Gettier counterexamples." In a Gettier counterexample, a person appears to justifiably, truly believe something without knowing it. The purpose of this paper is to outline a strategy whereby the proponent of TB can explain away Gettier counterexamples, or, to put it differently, the purpose is to sketch a believable story that TB's proponent can tell about why the appearance of true, non-knowing belief in Gettier counterexamples is *merely* appearance.

Let me define precisely what I take a Gettier counterexample to be. Edmund Gettier published "Is Justified True Belief Knowledge?" in 1963, and that paper instigated a certain well-known, generally recognizable dialectic among epistemologists that continues to the present day. The characteristic feature of that dialectic is the putting forth of an analysis of knowledge, typically "containing" justified true belief (i.e., analyzing knowledge as justified true belief or as justified true belief with some further characteristic), followed by the formulation of an apparent counterexample to the analysis. (Henceforth, I will sometimes

[1] This thesis has been defended by Crispin Sartwell in "Knowledge is Merely True Belief" and "Why Knowledge is Merely True Belief."

drop the qualifier "apparent" when doing so contributes to more natural expression.) Some of these counterexamples are counterexamples to the sufficiency of the proposed *analysans* for knowledge, some to the necessity. A natural way to define a Gettier counterexample would be as any of the counterexamples arising (in the past or future) from this recognizable dialectic. For my purposes, however, it will be useful to stipulate a narrower definition: A Gettier counterexample is any counterexample, arising from the dialectic just described, to the sufficiency for knowledge of an *analysans* "containing" justified true belief. This definition is historical, leaving open the question of the precise nature of the content of Gettier counterexamples. That question is sufficiently difficult that any answer to it should be argued for rather than prejudged in the definition; I will give my answer to the question and argue for it shortly.

I said above that I am going to sketch a believable story about why Gettier counterexamples merely appear to be counterexamples to TB. But a believable story is not necessarily one that ought to be believed. My story is, I will argue, one that very well could be true given everything else we know about the world, but it is just as likely to be false. If one is already inclined to accept TB, then one will have reason to accept my story. But almost all epistemologists are not so inclined, and so they can either take the story or leave it. It appears, then, that this paper will be of interest to only the narrowest of epistemological audiences. To remedy this worry, I would like to begin by briefly describing the motivation for holding TB, i.e., by explaining why I think one ought to be inclined to hold it. The ultimate argument for the position of this paper is that TB, together with believable stories explaining away both justification and Gettier counterexamples, provides the most theoretically satisfying, overall picture of knowledge we have available. My story explains away Gettier counterexamples and so provides an important component of this picture. To make the argument entirely convincing would exceed the scope of the paper; I only sketch it to make epistemologists aware of what I think has the potential to be a formidable position in logical space.

Two Reasons to Think Knowledge is Merely True Belief

Practically everyone accepts that true belief is necessary for knowledge. TB claims that it is sufficient as well. This is *not* to deny that other conditions may also be necessary. It is only to assert that, if there are further necessary conditions for knowledge, then those conditions are also necessary for true belief. A true analysis of knowledge might contain something in addition to true belief, but that additional element will be redundant. To see why proponents of TB assert this, let us focus on the traditional—though now defunct—analysis of knowledge, justified true belief (JTB). TB's proponent claims that the justification condition is redundant at best.

One reason to deny the necessity of specifying a justification condition in the analysis of knowledge (which is different from denying the necessity of justification for knowledge) is that it allows us to hold a simpler theory of what knowledge is. TB, quite obviously, is simpler than JTB. It is standard methodology to prefer a simpler theory to a more complex one, *ceteris paribus*. Thus, we should be naturally inclined to accept TB over JTB and, by extension, over all more complex analyses containing JTB.

The force of this reason for holding justification to be redundant at best is real, I think, but limited. After all, there is a parallel reason for holding truth to be redundant at best, or belief. The TB theorist is not motivated by considerations of simplicity alone; he feels that there is something particularly suspect about the inclusion of a justification condition in the analysis of knowledge beyond the mere additional complexity it adds to the theory. This feeling is the second reason for preferring TB to JTB. Let us examine the nature and source of the TB theorist's feeling about justification and then see whether a similar feeling about other possible conditions for knowledge arises from the same source.

The most prevalent conception of epistemic justification is of a property that a belief has in virtue of arising in the right sort of way. A belief is justified just in case it is produced by good reasons or good evidence. Some might quibble with this formulation, preferring to say that a belief is justified just in case it is produced by something like the activity of basing beliefs on good reasons or good evidence. The first way of speaking strikes me as perfectly natural, but I will not argue the point. In the sequel I will simply speak of justified beliefs as those produced by good evidence, but the reader may, if she wishes, substitute the longer formulation throughout. Nothing will turn on which formulation we adopt.

Now, when we ask what good evidence is, the best answer we are likely to devise is that it is something conducive to true belief. In some sense, good evidence consistently, reliably produces true beliefs. To put it another way, good evidence naturally has a high degree of power to produce true beliefs—this power in high degree is, in fact, the defining characteristic of good evidence. What is striking about the JTB analysis of knowledge is that it ascribes to good evidence another, distinct power: Not only does good evidence produce true beliefs, it also combines with (or interacts with) true beliefs (the very ones it has produced) to give rise to knowledge. These observations about good evidence are mirrored by observations about bad evidence (or the activity of basing beliefs on bad evidence) and non-evidence (or the activity of basing beliefs on things that are not evidence, or on nothing at all). Bad evidence and non-evidence are not conducive to true belief. They produce true beliefs only inconsistently and unreliably. They have no natural power to produce true beliefs, or at least have this power in a very low degree. And, though bad evidence and non-evidence do infrequently produce true beliefs, they have, on JTB, no power whatsoever to create knowledge through their interaction with those true beliefs. It is this double power attributed to good evidence, and this double deficiency of power attributed to bad evidence and non-evidence, that makes the justification condition of JTB seem fishy. We ought to be able to explain the intuitive lack of knowledge in cases of unjustified true belief, not by introducing a new power of good evidence lacking in bad evidence and non-evidence, but by appeal to the deficiency of the true-belief-producing power that is definitive of bad evidence and non-evidence. Let me make my meaning clearer by describing a pair of closely related thought experiments.

First, imagine that there is a clump of oak trees in my backyard. They regularly produce large numbers of acorns and drop them to the ground. Suppose that one oak tree in the clump is barren: it looks perfectly normal, but it never produces acorns. At least, it doesn't until one day when I walk by and notice a single acorn growing from one of its branches. Intrigued, I reach up to pluck the acorn off, but it crumbles to dust in my hand. In this situation, I think, I should be somewhat surprised, but not terribly surprised. For I already knew that the tree's acorn-producing power was defective, in that it rarely produced any acorns at all. When the tree does manage to produce an acorn, it is somewhat to be expected that that acorn will, in fact, be something less than the full kind of acorn produced by the other, normal oak trees. This tree's acorn is not a true acorn at all, but a mere acorn husk sharing no more than its superficial appearance with an actual acorn. Compare the situation of an unskilled baker trying to make a cake. He has no recipe, and no experience baking—he does not even know that cakes have to be cooked in the oven—but he merely tries to replicate the finished product he has seen and eaten in the past. We might imagine him combining various powders and liquids, trying to mold them into a cake shape and cover them with icing. Most of the time, what he produces does not even *appear* to be a cake—it is a strangely colored, amorphous blob or puddle. Yet, occasionally, the unskilled baker does produce something with the superficial appearance of a cake, though, upon cutting it and trying to eat it, we always discover it not to be a cake. The deficiency of the unskilled baker's cake-making power accounts both for the fact that he rarely makes anything that even looks like a cake, and for the fact that when he does make something resembling a cake, it merely resembles a cake. When performing at his very best the most that the unskilled baker can produce is the superficial appearance of a cake. Likewise, it is easy to suppose that the barren oak tree retains within itself the slightest flicker of the acorn-producing power, but even when that power operates at full capacity, the most it can produce is the superficial appearance of an acorn.

Now imagine (as a later episode in the first thought experiment) that I am passing by my oaks one day and notice that the shade of the barren tree's bark is slightly lighter than that of all the other trees. Just as the crumbling of the acorn husk surprised me somewhat, so too would the lightness of the bark. But I think the lightness of the bark would surprise me more. For I could not see it as a manifestation of that same deficiency—of the acorn-producing power—that I already knew the tree to have. Rather, the bark color would result from the deficiency of a distinct, bark-darkening power. No doubt I would postulate some common defect deep within the tree's genetic structure from which both of these deficiencies stem, but the point is that the discovery of this deep defect is far more surprising than that of another manifestation of the acorn-producing deficiency, of which deficiency I was antecedently aware.

Now for the second thought experiment. Again there is a clump of oaks, all of which are fruitful, except for one barren one. I have this time examined none of them closely, and I have never tried to pluck anything from the barren tree. I have noticed no variation in bark color. Looking out my back window, I see squirrels gather where the acorns fall from the fruitful trees. Where more acorns fall, more squirrels gather, and where there are fewer acorns, there are fewer squirrels. It would be perfectly rational for me to form the hypothesis that squirrels gather wherever acorns fall—that the presence of acorns is sufficient for the presence of squirrels. Suppose now that, very rarely, I see an acorn fall from the barren tree. The squirrels will not touch it. I now appear to have a counterexample to my squirrel-gathering hypothesis. In response, I can form one of two new hypotheses. First, I might hypothesize that the acorns from the barren tree are really not acorns at all—they merely superficially appear to be from the distance at which I view them. In fact, those "acorns" are but the defective products of a deficient acorn-producing power I already knew the barren tree to possess. Were I to examine the "acorns," I might hypothesize, I would find them to crumble, to be empty inside, or to be mere husks. The second hypothesis is that the fruitful trees possess, in addition to their acorn-producing power, a distinct power, e.g., a bark-darkening power. Here I hypothesize that it is in the nature of squirrels to be attracted, not to mere acorns, but to acorns near trees with dark bark. The barren tree, my hypothesis continues, is deficient in two ways: it lacks both the full acorn-producing power of the fruitful trees, and it lacks their bark-darkening power as well. From the first thought experiment, I think it is clear that, barring further investigation of the matter, the first hypothesis is the superior one and the one I ought to believe.

To avert any possible confusion, let us explicitly match up the elements of the second though experiment with their epistemological analogues. The fruitful oak trees are good evidence. The barren one is bad evidence. The acorns are true beliefs, and the gathering of squirrels is knowledge. The darkening of bark is good evidence's additional power (on the JTB theory) to combine with true belief to create knowledge. My simple, initial squirrel-attraction hypothesis (that acorns are sufficient to attract squirrels) is TB. The acorn from the barren tree that fails to attract squirrels is a justification counterexample, and the second of the two hypotheses formed in response to it is JTB. I claim that we have reason in epistemology to accept an analogue of the first of the two hypotheses over JTB. This analogue is that the "true beliefs" produced by bad evidence or non-evidence in justification counterexamples merely seem to be true beliefs. (They may actually not be true or not be beliefs, but I think it more likely that they are actually not beliefs.) When operating at its very best, the most that bad evidence can produce is something with the superficial appearance of a true belief. Now there is one major difference between our thought experiment and the epistemological issue of analyzing knowledge: Whereas it is very easy to see how some physical object could merely appear to be an acorn (it could be empty inside, or be disposed to crumble when touched), it is not so obvious how a psychological state could merely appear to be a true belief. We need a story of how we could be led to mistake something that is not a true belief for a true belief. Such a story would explain away justification counterexamples, and, I hold, the very fact that it explained them away and allowed us to adopt the analogue of the first hypothesis from the thought experiment would give us strong reason to accept it, even if no further argument for its truth were adduced.

This paper does not attempt to explain away justification counterexamples; it attempts to explain away Gettier counterexamples. Yet I believe that all the remarks I have made about the former task carry over to the latter. In a Gettier counterexample, a true belief is produced by a means that is not normally conducive to true beliefs. The true belief arises through luck or a freak accident. The source of the belief is deficient in its true-belief-producing power. As in the justification case, I hold that here we would do better to hold that the apparent true belief arising from the deficient source is no true belief at all, but merely has the superficial appearance of one, than to add another condition to the analysis of knowledge, granting to normal, true-belief-conducive situations an additional power to combine with true beliefs and create knowledge. But, again, we need a story of how the apparent true beliefs arising in Gettier counterexamples can merely appear to be true beliefs; and the very fact that such a story explains how this is possible gives us strong reason to accept it. After analyzing the content of Gettier counterexamples in more detail, I will tell such a story.

Gettier Counterexamples

I hypothesize that all Gettier counterexamples share a fundamental, common structure. I say "hypothesize" because I cannot individually consider every Gettier counterexample ever proposed and demonstrate the structure's presence in each, and thus I cannot definitively prove its universality. Instead, I will give six confirming instances—three in this section, three later in the paper—that span a fairly wide range among all Gettier counterexamples currently in existence. The story I plan to tell to explain away these six counterexamples will be targeted to the common structure, and so, I propose, we can reasonably expect that the story can be eventually expanded to explain away all Gettier counterexamples.

Briefly and abstractly, the common structure I have in mind is this: A subject encounters a certain object and comes to believe truly that it possesses a certain property. The object belongs to a larger class of similar objects, objects so similar that, had the subject encountered one of them in the same way as the object actually encountered, he or she would have come to believe—on the very same grounds—that the similar object possesses the property. Yet the vast majority of the objects in this larger class lack the property; had the subject encountered one of the other, similar objects, the resultant belief would almost certainly have been false. Let's observe this structure in a particular Gettier counterexample, the Barn Façades case:

> *Barn Façades*. S is driving through the countryside, looks up, sees a barn, and comes to believe that it is a barn. Thus, S truly believes that the thing he sees is a barn. However, there are in the vicinity a large number of barn façades, which are not barns but look indistinguishable from barns when viewed from the highway. S therefore does not know that the thing he sees is a barn (see Goldman).

The structure I have described is apparent in this case. In other cases it is harder to see. Take the following Gettier counterexample:

> *Disjunctive Addition*. S is given good evidence that Jones owns a Ford: Jones shows S a title to a Ford with his (Jones') name on it, he drives by in a Ford and offers S a ride, etc. Based on this evidence, S comes to believe that Jones owns a Ford. Now, S has a friend named Brown of whose whereabouts S is currently completely unaware. S uses the logical rule of disjunctive addition to infer from his belief that Jones owns a Ford the new belief that either Jones owns a Ford or Brown is in Barcelona. The twist is this: Jones does not own a Ford. The title was forged, the Ford he drove by in was rented, and so forth. But, by sheer chance, Brown is in Barcelona. Thus, S truly believes that either Jones owns a Ford or Brown is in Barcelona, but he does not know it (see Gettier).

The object encountered here, I claim, is a state of affairs, viz., Brown-being-in-Barcelona. S's encountering of it consists simply in his thinking it up randomly off the top of his head. The property S believes this state of affairs to have is being such that either Jones owns a Ford or it obtains. Believing that either Jones owns a Ford or Brown is in Barcelona is (the "is" of identity) believing that the state of affairs Brown-being-in-Barcelona is such that either Jones owns a Ford or it obtains. This state of affairs belongs to the class of all states of affairs, the vast majority of which are not such that either Jones owns a Ford or they obtain. For instance, the state of affairs Brown-being-in-Madrid lacks this property.[2]

I asserted above that believing that either Jones owns a Ford or Brown is in Barcelona is identical to believing that a certain state of affairs has a certain property. Let me pause now to justify this assertion. Suppose I am quizzing you on your beliefs. I ask you whether you believe that the sky is blue, that dolphins are mammals, that you're a brain in a vat, etc. After each question you introspect briefly and then reply with either a yes or a no. At some point in the sequence, I ask whether you believe that Brutus killed Caesar, and you say yes. Now imagine that the next question is "Do you believe that the killing of Caesar was perpetrated by Brutus?" It is likely that this question, falling where it does in the sequence, will initiate in you a thought process with quite a different phenomenology from the preceding ones. You will not simply introspect, find the belief, and answer in the affirmative. Rather, you will experience a feeling of recognition—recognition that the sentence embedded in this belief query says exactly the same thing as that embedded in the

[2] Strictly speaking, I am here talking not merely about states of affairs (all of which one might reasonably suppose to obtain, just as one may reasonably suppose that everything exists), but about possible states of affairs. Moreover, the large class of states of affairs I refer to must be taken to include only positive ones—it does not contain Brown-not-being-in-Barcelona in addition to Brown-being-in-Barcelona. If the class did contain negative as well as positive states of affairs, then, obviously, exactly half would obtain and half would not. Precisely how to distinguish positive from negative states of affairs is a difficult question, but it is not one I can enter into here.

immediately preceding one, only in more convoluted language. Alongside the recognition will be a feeling of repetition, a feeling that the present question simply repeats the preceding one, and that to say yes would simply be to repeat your last answer. Indeed, you might not say yes at all, but instead point out to me that I had merely asked you the same thing over again in different words. It is eminently plausible that what explains these feelings of recognition and repetition is that it is one and the same state—the same believing—that makes the answer to both questions affirmative. The believing that Brutus killed Caesar *is* the believing that the killing of Caesar was perpetrated by Brutus. If we now imagine ourselves quizzing S from the Disjunctive Addition case in this manner, asking him first whether he believes that either Jones owns a Ford or Brown is in Barcelona, then whether he believes that the state of affairs Brown-being-in-Barcelona is such that . . ., we can easily imagine that he would experience feelings of recognition and repetition like those just described. He would probably feel that the second question simply repeated the first in more convoluted language. Thus we may reasonably infer that the two believings underlying the two affirmative answers to our two questions are, in fact, one.

Let's see one more Gettier counterexample with the kind of structure I have been describing.

> *Stopped Clock.* There is a clock on S's wall that has stopped since the last time S looked at it, but S does not know that it has stopped. S looks at it, and it reads 5 o'clock, where it stopped. S thus forms the belief that it is 5 o'clock. By sheer coincidence, it actually is 5 o'clock when S looks at the clock, so S's belief is true. Yet S does not know that it is 5 o'clock (see Russell).

Here S encounters a moment, the present moment. He encounters it by focusing his attention on it while looking at the clock. The belief he forms about it is that it is 5 o'clock, i.e., that it has the property of being 5 o'clock. Believing that it's 5 o'clock is believing that the present moment has the property of being 5 o'clock. The present moment belongs to a larger class of moments, all those in the nearby temporal vicinity at which the hands of the clock read 5. Since the clock is stopped—has been for a while and will continue to be for a while--the vast majority of those moments are not 5 o'clock. Thus, had S encountered one of those other moments in the same way that he encounters the present one, he would have believed falsely that it was 5 o'clock.

Representative Concepts

I propose to explain away Gettier counterexamples by telling a story on which the subject does not believe the relevant proposition, viz., that the encountered object has the property in question. Instead, the subject believes a different, related proposition which is false: the proposition that all the objects in the larger class to which the encountered object belongs have the property. To tell the story, I will first need to draw two distinctions, and then to develop the apparatus of what I call representative concepts.

The first distinction is the familiar one between sentences and propositions. Sentences express propositions, and often different sentences express one and the same proposition. "It's raining today" uttered on Monday and "It rained yesterday" uttered on Tuesday express the same proposition, namely, that it rains on Monday. The second distinction is between believings and believeds. The term "belief" exhibits an –ing/–ed ambiguity. Believings are psychological states of subjects. Believeds are mind-independent entities—propositions. We must be careful not to confuse these two things.

Now the notion of a representative concept is key to my explanation of Gettier counterexamples, so I will devote significant space to its explication. First, I need to say a few words about what I mean by the more general term "concept." I don't mean anything very specific. I take a concept to be any kind of non-perceptual, non-sensory mental representation. Roughly, my concepts correspond to Hume's ideas, as opposed to his impressions. In fact, I do not believe that the term "idea" is any less appropriate for what I have in mind than "concept." We think thoughts about the world, and those thoughts contain elements representing things in the world, and these elements are what I call concepts. Concepts are constituents of believings. (I leave it open whether there are corresponding constituents of believeds, i.e., propositions.)

I make two important assumptions about concepts. First, I intend for the notion of concept to be wide enough to include particular or individual concepts—concepts that represent one thing only and are essentially not multiply applicable. This strikes me as no abuse of the language. It is perfectly natural to speak of my concept of Julius Caesar or of London. Still, if the only kind of concepts you recognize are general, multiply-applicable concepts, then I am using the term "concept" more broadly than you. When I think about Julius Caesar, there is some constituent of my thought that represents or stands for Julius Caesar, and that constituent, whatever you may prefer to call it, I call the individual concept of Julius Caesar.

The second assumption I make about concepts is that they have or contain modes of presentation. A concept of something represents it *as* something; at least, concepts commonly do this, if not always. This fact seems pretty obvious to me. My concept of kittens represents them *as* baby cats. My concept of Julius Caesar represents him *as* an ancient Roman general. Another way to put this same point is that concepts in some sense contain information. There are some facts about kittens and Julius Caesar that I would naturally describe as part of my very concept of those things. It is part of my concept of kittens that they are baby cats. If you told me kittens were not baby cats, I would probably insist that you were using the word "kitten" differently from me. Or, if I were to accept your claim with the sense it straightforwardly appears to have, then I would have to do something drastic, such as "fundamentally revise my concept of kittens," or "reject my current kitten concept in favor of a new one," whatever these phrases might mean. By contrast, it is no part of my kitten concept that, on average, kittens take a certain number of days to open their eyes. I could accept or reject this fact without disturbing my kitten concept in any fundamental way. I acknowledge that talk of concepts "containing" information is unsettlingly metaphorical, and I allow that whether or not a certain concept contains a certain piece of information may be a matter of degree and/or vague. All I claim is that there is some point to this way of speaking—some truth about the nature of concepts that it at least roughly captures.

Now I am in a position to introduce the idea of a representative concept. We must first be careful to distinguish representativeness from representationality. All concepts are representational, because they all represent something, but not all concepts are representative. As I will use the term, only individual concepts can be representative. I should also mention that I have a second name for representative concepts, "arbitrary" concepts. Though I believe that the term "representative" better captures the nature of these concepts in most cases, "arbitrary" will occasionally be more appropriate. Rather than give a precise definition of representative/arbitrary concepts right up front, I will introduce them by way of examples, describing each example in relatively loose, inexact terms in order to fix the basic idea in the reader's mind. I will make my account more exact after considering these examples.

I shall begin with the example that first inspired in me the idea of representative concepts, though, admittedly, it is probably the least straightforward of the examples I will consider. I will first describe it in a way that is natural and intuitive, but strictly incorrect, and then redescribe it correctly. The example is a common practice used by mathematicians in constructing proofs. Suppose a mathematician is trying to prove that all members of a certain class, C, have a certain property, Φ. The mathematician will likely begin with a statement such as this: "Let x be an arbitrary member of C." Then she will go on to deduce certain statements about x, say, "x is φ_1," "x is φ_2," . . ., and "x is φ_n," relying only on the fact that x is in C. Finally, the mathematician will conclude "x is Φ." This routine is taken to prove the statement "All members of C are Φ." The statements of this proof, starting with "x is φ_1" and ending with "x is Φ," are all about a single object, x. Corresponding to these statements are thoughts in the mathematician's head, and in the heads of anyone who may read her proof, about the single object x. And yet, in some crucial sense, these statements and thoughts about x are *also* statements and thoughts about every member of C; that is why they amount to a proof of a conclusion about the entire class. The statements and thoughts in question seem to play two roles, being both singular and general at the same time. More specifically, the concept of x appearing in the thoughts appears to be both an individual concept—of x—and a general concept—of the members of C. The dual nature of the concept of x seems to be captured by the word "arbitrary" appearing in the first sentence of the mathematician's proof. This word indicates that there is nothing special or distinctive about x among the members of C; x is a perfectly ordinary, generic member of C, normal in every way. It is as if the mathematician stuck her hand into C and pulled out one of its members completely at random to work with in her proof. When she goes on to prove statements about that member, we see that she could just as easily have picked any other member of C at the beginning, and the same statements would have held for that other member. It is because we conceive of x as arbitrary—because arbitrariness is part of our concept of x—that when we think that it has the property Φ, we *thereby* think that all members of C have Φ. Another way to describe the situation at hand is to say that our concept of x is a concept of it as *representative* of all the members of C (or simply of C). We can narrow this last description a bit by saying that we conceive of x as representative of C with respect to the property Φ. There is nothing in our concept of x that has any bearing, so far as we can tell, on whether it is more or less likely than any other member of C to have Φ; any information in the concept of x relevant to the having or lacking of Φ is equally a part of our concept of any other member of C as a member of C. Of course, this relativization of the representativeness of the concept of x to the property Φ is, in the case at hand, needless. We conceive of x as representative of C with respect to any property whatsoever. Nonetheless, in the examples about to come, this kind of relativization will be essential.

I said above that my initial description of the mathematical example would be strictly incorrect, and it was. I pretended there that the class C contained a distinct object named "x" and that we had an individual concept of it. But surely this was pure pretense. The symbol "x" no more names one member of C than any other, and when we "think about x" we no more think of one member than any other. To think about a particular object, it seems to me, one must conceive of that object as in some way distinguished from all others. To think about a particular toothpick from a box of 1,000 qualitatively identical ones, I must at least hold it up before my eyes and conceive of it as "the one presently before my eyes." If the mathematician in our example could literally reach into C and pull out a particular member, then we could think of that member—conceived of as the one in her hand—that it was Φ. But she can only figuratively reach in, so we cannot think any such thought. My remarks about the mathematician's proof are still apt, but they must be reformulated in terms, not of x, C, and Φ, but of our symbols "x," "C," and "Φ." Here is the correct description of the situation:

The syntax of the mathematician's language specifies a certain subset of that language's symbols as "designators"—symbols whose business it is to refer to objects. Designators are contrasted with symbols like logical connectives and quantifiers, which are in a different business. The symbol "x" is a designator. When the mathematician writes the proof described above, she is demonstrating—or perhaps I should say displaying—that "x" has a certain syntactic property Y: a proof can be written with a single premise comprising it—the symbol with Y—followed by the string "is a member of C" and with a conclusion comprising it followed by "is Φ." (The notion of proof is, of course, purely syntactic.) When we read the mathematician's proof, we are led to believe that "x" has Y. Moreover, we see that there is nothing special or distinctive about "x" among the class of all designators with respect to having or lacking Y. The mathematician could just as easily have chosen "y," "z," "w," or any other designator and demonstrated, in precisely the same way, that that other designator has Y. Our concept of "x" is of it as an arbitrary or representative designator with respect to Y. Thus, when we think that "x" has Y, we thereby also think that all designators have Y. Now, we know that when all the premises of a proof are true, its conclusion must also be true. We also know that a sentence comprising a designator designating a member of

C followed by the string "is a member of C" is true. Since every designator designating a member of C has Y, it follows that, for every such designator, a proof can be written with all true premises whose conclusion comprises that designator followed by "is Φ." Hence, every sentence comprising a designator designating a member of C followed by "is Φ" is true. We therefore conclude that every member of C is Φ.

In this description of our mathematical example, the notion of a representative concept still plays an essential role. This time, however, the concept is of a symbol as representative of a class of symbols with regard to a syntactic property. By thinking that that symbol has that property, we think that all symbols in the class have the property. We then go on to make further inferences concerning designators of members of C and concerning members of C themselves, but these inferences, though mathematically important, are not germane to the topic of this paper. It is important to note that, in the correctly described example, "x" is conceived of as representative of all designators with respect to the property Y, and now the relativization to a property is essential. The symbol "x" is not representative of the class of all designators *tout court*, because it is distinguished from all the others by its shape. This distinction is what allows us to have an individual concept of it in the first place. If we were to think that "x" has some very specific shape property S, then we would not thereby think that all designators have S. The shape information built into our concept of "x" bears on whether or not "x" *in particular* has S. Since that information has no bearing on whether "x" has the syntactic property Y, however, we can think that all designators have Y by thinking that "x" has Y.

Let us turn to a different example. Suppose you are watching a show about animals on television. At one point the host of the show holds up a strange-looking creature whom he introduces as “Lenny.” “Lenny,” the host says, “is a lemur.” As the camera zooms in on Lenny, the host describes some of the animal’s characteristics. He tells you that Lenny has a certain kind of diet, has a certain lifespan, lives in a certain climate, and makes certain sounds to attract mates. The host says all these things about Lenny, but it is clear that, in some sense, he is also saying them about all lemurs. Furthermore, you are both thinking about Lenny, but in some sense, you are also thinking about all lemurs. Your concept of Lenny seems to be playing two roles, both individual and general at the same time. As far as you can tell, there is nothing special or distinctive about Lenny, among the class of all lemurs, with respect to his diet. Nothing in your Lenny concept has any bearing on whether or not he *specifically* has the kind of diet described by the host. You conceive of Lenny as representative of all lemurs with respect to having that diet, and so to think that he has it is also to think that all lemurs have it. By contrast, if the host said that Lenny had a disease that required his keepers to feed him a special diet, you would not, in thinking this thought, be thinking that all lemurs have the special diet. As you think the thoughts expressed by the host, you look at Lenny and have him in particular before your mind. But your concept or idea of him seems to be serving not so much as the subject of singular thoughts (in the sense of “subject” that is contrasted with “predicate”), but as a psychological aid to help make concrete and vivid a number of general thoughts about all lemurs.

Two more quick examples. First, a physicist is having dinner with his family and, in the course of conversation, describes something as composed of atoms. Curious, his young son asks him which things are made of atoms. The physicist replies, “Everything in the world is made up of atoms. That fork you’re holding right now—that’s just a big collection of atoms.” The son conceives of his fork as representative of all physical objects with respect to atomic composition. By believing his fork to be made of atoms, he believes all physical objects to be made of atoms. Second, you are walking down the street with a friend who has recently become delusional. Your friend believes that a worldwide conspiracy is working toward his demise. “Everybody is in on it—everybody!” he swears to you. At that moment he turns and points to a stranger on a distant street corner whom you know neither of you has ever seen before. “That guy is a part of it! That guy over there is trying to get me!” he says. Both you and your friend conceive of the stranger as representative of the class of all people with respect to membership in a worldwide conspiracy. You think that the stranger is part of the conspiracy and thereby think that everyone is. (More specifically, your friend believes it and you simply entertain it.) In both of these cases, something is conceived of representatively after being selected in a particularly random, unprincipled way. It is clear that the son could have asked his question in any context, and his father could have pointed to any salient physical object in that context. Likewise, your delusional friend could have pointed to anyone in any situation in which you had had your conversation with him. The randomness with which the objects of discussion in these examples are selected emphasizes that there is nothing special or distinguished about them. Whatever goes for them goes for every member of the relevant class to which they belong.

With these examples under our belt, let us try to describe more exactly the nature of representative concepts; in the next section I will try to spell out precisely their connection with beliefs. First, it should be clear that it is too simple to think of a concept as representative *tout court*. Representativeness is a doubly relational property. A representative concept is always a concept of an object as representative of a class C with respect to a property Φ. What it means for your concept of an object to be of it as representative of C with respect to Φ is for there to be, insofar as you conceive of it, nothing special or distinctive about that object, among all members of C conceived of as members of C, relevant to the having or lacking of Φ. No part of the information in your concept has any bearing on the question of whether the object has Φ, at least, no part of the information which is not also part of the information in your concept of any member of C as a member of C. Most individual concepts are probably of their referents as representative of many classes with respect to many properties, and as not representative of many more classes with respect to many more properties. Your concept of Lenny, to refer back to an earlier example, is of him as representative of the class of lemurs with respect to having a certain diet, but perhaps also as representative of the class of mammals with respect to being warm-blooded. If he has an injury, however, you do not conceive of him as representative of all lemurs with respect to being injured.

This last observation about representativeness may give rise to an objection that my picture builds an unrealistic amount of information into our concepts. You are supposed to conceive of Lenny as representative of lemurs with respect to one property, of mammals with respect to another, and of many other classes with respect to many other properties. But surely your concept of Lenny does not contain this glut of information; you are thinking far less than all of this when you think of Lenny. I reply that representativeness is not a matter of a concept containing information; on the contrary, it a matter of a concept *lacking* certain information. Representativeness is a negative or privative property of a concept. You conceive of something as representative of a class with respect to a property when there is nothing in your concept with any bearing on whether or not the thing has the property, at least, nothing that is not equally contained in your concept of any other member of the class as a member of the class. Thus, even a relatively simple concept can be of something as representative of many classes with respect to many properties.

Representative Concepts and Belief

I now propose to describe, as precisely as possible, the nature of the connection between representative concepts and beliefs. Let me first remind the reader that my task here is merely to *explain* how Gettier counterexamples can appear to be counterexamples to TB without being so in fact. As such, I will not directly argue for many of the assertions I make in this section. The ultimate argument for the view presented in this section, it will be recalled, is that it, together with TB and a (still forthcoming) story explaining away justification counterexamples, provides the most theoretically satisfactory overall picture of knowledge available to us. I acknowledge up front that some elements of my view will be, to some extent, theoretically *un*satisfactory. It makes the nature of beliefs and the semantics of belief ascriptions more complicated than we previously thought them to be. I will try to mitigate this problem as much as possible by giving examples to show that the complications I introduce are closely analogous to complexities we do have independent reason to believe exist. Thus, even though my view introduces new complexities into the theoretical picture of belief, it at least does not introduce any radically new *kind* of complexity. I primarily want to stress that the reader must not balk when I make certain substantive, non-obvious claims without any positive argument. Some of them will appear to be, on their face, positively *counter*intuitive, and where they do so I will argue that they are not. Everything I am going to say is, I believe, at least intuitively neutral—our intuitions weigh neither for or against them. More than this neutrality, however, I cannot always show.

In the previous section we saw cases in which a subject S conceived of an object as representative of a class C with respect to a property Φ, and which we described loosely as cases in which S believed that the object was Φ but, because of his recognition that the object was in no way special or distinct (among the members of C with respect to the having of Φ), also thereby believed that all members of C were Φ. There are two propositions in play in a case such as this: first, the singular proposition P_s, which we might express with the sentence "The object is Φ"; and second, the general proposition P_g, which we might express with the sentence "All members of C are Φ." Now I assert the following two claims about the cases under consideration: S believes P_g, and S does not believe P_s. The second of these two claims follows from a stronger one, which I will give a special name:

> **The Inability Thesis**: If S conceives of an object as representative of a class C with respect to a property Φ, then S cannot, so long as he or she conceives of the object in this way, believe the singular proposition normally expressed by the sentence "The object is Φ," *i.e.*, P_s.

The Inability Thesis seems wildly implausible. Let me give a reason to think that it is not: The most straightforward way to deny that someone can hold a particular belief (= believing) is to deny that he or she possesses the concepts essential to the belief. I want to suggest another, closely related way to deny someone the ability to hold a belief: acknowledge that he or she possesses the essential concepts and can hold some beliefs involving them, but assert that those concepts are too thin or informationally impoverished in a certain way to allow him or her to hold the particular belief in question. Let's look at an example where we might make just this type of denial.

Imagine that a young boy has learned the concept "happy" through exposure to a certain class of examples of happiness. Every case of happiness he has ever seen and heard described as a case of happiness has involved someone laughing, giggling, grinning, jumping up and down with giddiness, or something like that. Now, the boy's father is a soldier who, a few years ago, went missing in action and is presumed dead. One day, the boy's mother gets a call from the army saying that her husband has been found alive and will be returning home soon. The mother bursts into tears and, when her son asks her why she is crying, she tells him that it is because she is extremely happy. It is almost certain that the boy would be deeply puzzled by this claim of his mother's. The reason for his puzzlement, it is plausible to think, is that his current concept of happiness does not allow him to believe that his mother, in her current state, is happy. He cannot even entertain the thought; his mother's utterance does not express a proposition on which he can get any cognitive grasp whatsoever. The boy's concept of his mother in her current state, on the one hand, and of happiness, on the other hand, simply will not "link up" in the right way to form a belief (= believing). (His mother's happiness in her current state is, for him, a colorless green idea sleeping furiously.) The reason for the boy's cognitive inability is that his concept of happiness is too impoverished; it contains too limited a range of information, leaving out an entire dimension of the property of happiness. There are ways of being happy very different from any the boy has ever encountered and that form no part of his current concept. Since his mother is currently happy in one of those unfamiliar ways, her son is simply unable to think of her current state as a happy one. Of course, the boy may revise and enrich his happiness concept in such a way as to take account of his mother's state, and then he would be able to believe her happy; cases like the one at hand are just the kind that tend to initiate conceptual revision. But, as long as the boy's happiness concept remains unchanged, he cannot think that his mother is currently happy. Now it does not seem right to *deny* that the boy in this example has the concept (or, at least, a concept) of happiness at all, or even to deny that he has the "correct" concept. In some important sense, the boy's happiness concept is "the same as" your and my happiness concept. The only difference is that the boy's concept is thinner than ours; it contains less information, but enough important information to make it count as a concept of happiness nonetheless.

This example suggests that, when a concept belonging to a person is informationally impoverished, or lacking in content in a certain way, then that person may not be able to hold certain beliefs (= believings) involving it, even though he or she can hold others. Now, I said above that what it means for a concept to be representative is for it to be lacking a certain sort of information, viz., information to the effect that its object is special or distinguished among the members of a certain class with respect to the possession of a certain property. In light of this fact, along with our example, the Inability Thesis does not look so wildly implausible as it did at first.

There are probably two more sources of intuitive opposition to the Inability Thesis lurking in the reader's mind. The first is the straightforward thought that we do ascribe to people singular beliefs in cases involving representative concepts. Of the person watching the show about Lenny the lemur, we would naturally and, it seems, truly say that he believes that Lenny has such-and-such a diet. I neutralize this objection by resorting to some semantic trickery. The belief ascription, I acknowledge, is true, but what makes it true is not that the subject believes the singular proposition normally expressed by the sentence "Lenny has such-and-such a diet." Without being at all specific, I assert that sometimes a belief ascription involving a sentential that-clause can be made true by the subject's believing a proposition that is strictly logically stronger than the proposition normally expressed by the embedded sentence. To be more specific, I assert that sometimes (but perhaps not all the time) a belief ascription of the form “S believes that x is Φ” can be made true by the fact that S believes the conjunctive proposition normally expressed by the sentence “All Ψs are Φ and x is a Ψ.” So what makes true the ascription “S believes that Lenny has such-and-such a diet” is the fact that S believes the conjunctive proposition normally expressed by the sentence “All lemurs have such-and-such a diet and Lenny is a lemur” (the second conjunct being a conceptual truth for S, since the concept of lemur is contained in the concept of Lenny). This proposition implies, but is not implied by, the proposition normally expressed as “Lenny has such-and-such a diet.”

Here is an example to lend some plausibility to my claim about the semantics of belief ascriptions: A dictator comes to power in a foreign land containing a certain ethnic group, the Eths, and institutes a policy of discrimination and oppression against them. In particular, he forces all Ethish children to attend substandard schools. One day, an Ethish boy named Joe asks his father why he must attend the worst school in town. The father replies to Joe, “It’s because our leader believes that you are inferior.” This strikes me as a perfectly normal way a real person might talk, and I would not hesitate to describe his belief ascription as true. Yet the proposition normally expressed by its embedded sentence is the singular one that Joe is inferior, and we can legitimately suppose that the dictator does not believe this proposition because he has never encountered Joe and has no individual concept of the boy at all. In this case, of course, the dictator does *not* believe the proposition normally expressed as “All Eths are inferior and Joe is an Eth.” He believes only the first conjunct—only the major premise of the syllogism, as it were. But it seems plausible to suppose that what makes the father’s belief ascription true is the fact that the dictator believes the first conjunct, together with the fact that the dictator could easily find out the second conjunct stating that Joe is an Eth—by, say, looking at Joe or reading certain government records on Joe. This example is not perfectly analogous to the Lenny case, but I think it is close enough to lend some credibility to my assertions about the semantics of belief ascriptions in the latter case and, by extension, in all other cases involving representative concepts.

There is one more source of intuitive aversion to the Inability Thesis. It is the feeling that, in denying that we can believe certain singular propositions in cases where we have traditionally thought that we could, it is somehow depriving us of something. We like our beliefs, and we naturally resist anything that threatens to take them away. When made explicit, it is clear that this objection to the Inability Thesis has no force. If the Thesis is true then it has always been true, and to assent to it is no more than to recognize that we lack something we never had to begin with. Nevertheless, perhaps I can assuage the feeling of deprivation by suggesting that it stems from a certain picture of the nature of belief (= believing) that there is no good reason to accept. It is the picture of believings as discrete, mutually exclusive, non-overlapping objects, like sentences listed one after another on a page. This picture may be correct (whatever correctness could consist in here), but there is another one that is just as likely to be correct.

Consider clay statues. A clay statue is a statue-shaped piece of clay, but it cannot be merely that. For it contains within it innumerable other statue-shaped pieces of clay that we would not want to call clay statues, simply for the reason that they are proper parts of another piece of clay, viz., it. To be a clay statue is, to a first approximation, to be a statue-shaped piece of clay that is not a proper part of any other piece of clay. The property of being a clay statue is, in part, an extrinsic property. Now to have a clay statue is to have all the statue-shaped pieces of clay it contains, only in such a form that they lack the extrinsic property which would make them statues. You do lack those pieces *as statues* in one way, but not in the way that you would lack them as statues if someone stole the whole statue that they compose. As far as their intrinsic constitution is concerned, you possess them to the same degree whether they compose the larger statue or whether they are carved out and made into statues of their own.

Perhaps believings are, at least sometimes, more like clay statues than like sentences. Perhaps there are informationally contentful mental state—belief-"shaped" pieces of information, so to speak—only some of which count as beliefs because the property of being a belief is partially extrinsic. To be a belief is to be a belief-"shaped" piece of information that bears a certain relation to other belief-"shaped" pieces of information: it is not a "proper part" of any other belief-"shaped" piece of information. (The scare quotes here are, of course, all crucial.) To have a belief is to have all the belief-"shaped" pieces of information it contains, only in such a form that they lack the extrinsic property which would make them beliefs. You do lack those pieces of information *as beliefs* in one way, but not in the same way that you would lack them if you completely forgot the whole belief that they "compose." As far as their intrinsic constitution—and, in particular, their informational content—is concerned, you possess them to the same degree whether they "compose" the larger belief or stand on their own as beliefs.

The Inability Thesis denies one the ability to believe certain singular propositions (e.g., the one normally expressed as "Lenny has such-and-such a diet"), but it does so only in cases where one can believe a strictly logically stronger proposition (e.g., that all lemurs have such-and-such a diet and Lenny is lemur). In logic we often speak of the premises of a deductively valid argument as containing all the information in the conclusion. I want to suggest that someone who believes the proposition normally expressed as "All lemurs have such-and-such a diet and Lenny is a lemur" has a belief (= believing) containing the information that Lenny has such-and-such a diet. Though he has no believing of the proposition normally expressed as "Lenny has such-and-such a diet," he has a belief-"shaped" piece of information that is intrinsically identical to such a believing and would count as such a believing if only it were not "contained" in the believing of the logically stronger proposition. If we recognize that this way of viewing the beliefs in the Lenny scenario is no less valid than the normal way (i.e., as sentences listed on a mental "page"), then the fear that acceptance of the Inability Thesis will deprive us of something valuable should, at least largely, subside.

Let us take stock. I have asserted that in cases like those described in the preceding section, in which a subject holds a belief involving a representative concept, the subject does not—and cannot—believe the singular proposition we might naturally take him or her to believe. I have argued that this assertion, if not positively intuitive, is at least not counterintuitive. Further, I have asserted that in the type of scenario under discussion the subject does believe a certain general proposition, e.g., that all lemurs have such-and-such a diet.

Now it will be recalled that when we first introduced this type of scenario, we made statements to the effect that a subject S believes that an object x (conceived of as representative of a class C with respect to a property Φ) is Φ and *thereby* believes that all members of C are Φ. We have now seen that S's believing that x is Φ consists only in S's believing the proposition normally expressed as "All members of C are Φ and x is a member of C." Yet the word "thereby" was central to our description, and it strongly suggests, at the very least, that there is some important connection between the individual concept of x and the believing of the general proposition that all members of C are Φ—that some thinking of x, and of x specifically, somehow supports or underwrites the believing of the general proposition. What can this connection between the individual concept and the believing of the general proposition be if the concept forms no part of the believing?

The connection is a purely psychological one: Entertaining the concept—by which I mean brining it to the forefront of one's mind and using it to attend to or concentrate on its referent—serves as a psychological aid to holding the belief. The human mind is feeble, its powers limited. It is most comfortable thinking of the concrete, the particular, the fully determinate—this chair, that tree, that lemur. Only with great difficulty can the mind grasp hold of things abstract, general, and indeterminate. When it must think generally, it often brings to its aid the thought of something particular, a fully determinate, concrete individual falling under the abstract concept that figures in the general thought. The concept of this individual serves to support the abstract thought without actually figuring into it. We might say, without too much abuse of the language, that the individual concept is playing an important psychological role with respect to the belief, but it is not playing any doxastic role.

An analogy might be drawn between the use of individual concepts I am proposing and the use of a mnemonic device. I believe that the biological taxonomy of organisms is kingdom, phylum, class, order, family, genus, species. I am able to believe this, at least in part, because I have the concept of the sentence "Kings play chess on fuzzy green couches." Whenever I am called on to state the biological taxonomy, I always begin by bringing (the concept of) that sentence to mind. (Not the concept of the state of affairs represented by the sentence, mind you—just the sentence itself as a linguistic object.) If I forgot the sentence or lost my concept of it, I would, quite likely, be unable ever to recall the biological taxonomy. I dare say I would lose the belief that the taxonomy is kingdom, etc. altogether. My concept of the sentence "Kings play chess on fuzzy green couches" supports my belief about the biological taxonomy even though that concept clearly does not enter into the belief. Its role is purely psychological, not at all doxastic. I suggest that representative concepts can support general beliefs in a closely, if not perfectly, analogous way.

Gettier Counterexamples, Again

I will now apply the apparatus of representative concepts, developed in the last two sections, to six Gettier counterexamples—the three already presented and three others. My purpose in each case is to explain how the counterexample can merely appear to be an instance of true, non-knowing belief.

We begin with Barn Façades. Let us coin the predicate "apparent barn" to apply to all and only those things which are either actual barns or barn façades. Let us also suppose, just for ease of discussion, that the apparent barns in the vicinity of the actual barn S sees are numbered—from 1 to 100, say; and suppose that S sees #47, which is the only actual barn. Now, my explanation is this: S sees #47 and comes to believe that #47 is an actual barn. But his concept of #47 is of it as representative of all apparent barns (in the vicinity, at least) with respect to the property of being an actual barn. There is nothing in his concept of #47 with any bearing on its actual barnhood that is not equally a part of his concept of any other apparent barn as an apparent barn. Thus, the proposition he believes is the one we would normally express as "All apparent barns (in the vicinity) are actual barns." This proposition is false, so S does not know it. The proposition we would normally express as "Apparent barn #47 is an actual barn" is true, but, by the Inability Thesis, S does not and cannot believe it. The reason the Barn Façades case appears to be a counterexample to TB is that we fail to distinguish between sentences and the propositions they express. In describing the case, we truly make both a truth ascription—"It is true that #47 is a barn"—and a belief ascription—"S believes that #47 is a barn." Since the same sentence appears in both ascriptions, we assume that one and the same proposition is both true and believed. But that is false. The semantics of belief ascriptions is messy; they sometimes ascribe belief in propositions other than the ones normally expressed by the sentences appearing in them. This semantic deviance occurs in a number of different kinds of case, including those that involve representative concepts.

I explain the Disjunctive Addition counterexamples as follows: S believes that Brown-being-in-Barcelona is such that either Jones owns a Ford or it obtains. But he conceives of Brown-being-in-Barcelona as representative of the class of all states of affairs with respect to the property of being such that either Jones owns a Ford or it obtains. Thus, the proposition S believes is the one we would normally express by saying "Every state of affairs is such that either Jones owns a Ford or it obtains." This is false of course. The state of affairs Brown-being-in-Madrid is not such that either Jones owns a Ford or it obtains. Another way of describing S's believing is by saying that he believes that either Jones owns a Ford or an *arbitrary* state of affairs obtains, and this description seems exactly right. As the case is described, S merely pulled the sentence "Brown is in Barcelona" randomly from thin air and disjoined it with "Jones owns a Ford." Brown-being-in-Barcelona appears to be, in

S's mind, a mere stand in for any state of affairs whatsoever. It should now be clear how the rest of the explanation of the Barn Façades case carries over to the present one.

An objection may be raised to my explanation of the Disjunctive Addition case, an objection whose reply is already contained in the explanation. The objection is that I hold S's belief (in the sense of either believing or believed—the objection could be raised for each) to have the logical form of a universally quantified statement, whereas it surely has the form of a *disjunction*. My reply is that, although S's belief is not disjunctive, it nonetheless contains an important disjunctive part, viz., what follows the "such that" when I say S believes that every state of affairs is such that. . . . The presence and importance of this part accounts, I claim, for the intuition that the belief is disjunctive. In other cases we will also find that S's belief is *prima facie* of a certain logical form but that this form adheres only to one essential component of the actual belief, which is always a universally quantified statement.

The explanation of the Stopped Clock case runs as follows: Believing that it is 5 o'clock is believing that the present moment m is 5 o'clock, *i.e.*, has the property of being 5 o'clock. In the case at hand, let us suppose that S is looking at the clock at m, which is 5 o'clock, and believing at that moment that it is 5 o'clock. S has an individual concept of the moment m. How do we conceive of a moment? What information goes into our idea of it? A natural answer is that we conceive of it in terms of what we perceive to be happening (or saliently happening) or what state of affairs (or salient state of affairs) we perceive to obtain at the moment. The most salient state of affairs S perceives to obtain at m is that the hands of the clock read 5 o'clock. So, S conceives of m as a moment at which the hands of the clock read 5 o'clock. I claim that S conceives of m as representative of the class of moments at which the hands of the clock read 5 o'clock. From S's point of view, there is nothing special or distinguished about m, among all moment at which the hands read 5 o'clock, with respect to the possession of 5 o'clock-hood. No information in S's concept of m bears on the question of whether or not m is 5 o'clock that is not equally a part of his concept of any moment in which the hands read 5 o'clock, conceived of as a moment in which the hands read 5 o'clock. Thus, S's believing that m is 5 o'clock is equivalent to his believing that all moments at which the hands read 5 o'clock are 5 o'clock. This latter proposition is false, for by stipulation the clock is stopped and so has recently read, and will soon read, 5 o'clock at moments which are not 5 o'clock.[3]

Here is a Gettier counterexample we have yet to consider:

> *Existential Generalization*. S knows that there are two people in his office, Nogot and Havit. Nogot has given S a lot of good evidence to believe that he, Nogot, owns a Ford. Havit has given S no such evidence. S forms the belief that Nogot owns a Ford. Then, S uses existential generalization to form the belief that someone in his office owns a Ford. Now, it turns out that Nogot actually has no Ford, while Havit does have one. So S's belief that someone in his office owns a Ford is true, but he does not know that someone in his office owns a Ford (see Lehrer).

To explain the Existential Generalization case, let's coin a new predicate, "Nogot-set." A Nogot-set is just any set containing Nogot. Believing that someone in S's office owns a Ford is believing that the Nogot-set {Nogot, Havit} is such that someone in it owns a Ford. S conceives of {Nogot, Havit} as representative of all Nogot-sets with respect to the property of being such that someone in it owns a Ford. And thus, in believing that {Nogot, Havit} is such that someone in it owns a Ford, he believes the proposition we would normally express

[3] To make the explanation of this case more plausible, we might restrict the class of moments in question to moments at which the hands read 5 o'clock *in the nearby temporal vicinity*, i.e., during the previous few weeks or months and during the upcoming few weeks or months. The reason for this restriction is that S may conceive of moments when the hands read 5 o'clock but which are also hundreds of years in the future when the clock will certainly no longer work. The information in *those* individual concepts may very well fail to bear on the question of whether the moments they refer to are 5 o'clock. Making a restriction like this does not, so far as I can tell, vitiate my account.

by saying that every Nogot-set is such that someone in it owns a Ford. This is false, since no one in {Nogot} owns a Ford. It is clear that, in performing existential generalization, S picked the Nogot-set {Nogot, Havit} randomly; he could just as easily have believed that someone in the building containing his office, or in the state containing the building, or on the planet earth owns a Ford. S believes that an arbitrary Nogot-set is such that someone in it owns a Ford.

The first counterexample from Gettier's original paper is the following.

> *Definite Description*. Two people have interviewed for a job, S and Jones. S has counted the coins in Jones' pocket and knows that there are exactly ten. S also has very good reason to believe that Jones will get the job. S comes to form the belief that the person who will get the job has exactly ten coins in his pocket. Yet, unbeknownst to S, it is he who will get the job, and he also happens to have exactly ten coins in his pocket. S truly believes that the person who will get the job has exactly ten coins in his pocket, be he does not know it.

The obvious first question to ask about this case is how we are supposed to interpret the definite description appearing in the belief ascription—a thorny issue in the philosophy of language. One traditional interpretation is the Russellian one: Believing that the person who will get the job has exactly ten coins in his pocket is believing that there is someone who will get the job and who has exactly ten coins in his pocket. The sentence contained in this latter description of the believing is an existentially quantified statement, and we have already seen how to handle those. First we bring out its domain of discourse, which is presently implicit: Believing that there is someone who will get the job and who has exactly ten coins in his pocket is believing that there is someone in the set {Jones, S} who will get the job and who has exactly ten coins in his pocket. This is equivalent to believing that the set {Jones, S} is such that someone in it will get the job and has exactly ten coins in his pocket. The set {Jones, S} is a Jones-set, and an arbitrary one in this case, so S really believes that every

Jones-set is such that someone in it will get the job and has exactly ten coins in his pocket. This belief is false, because no one in the Jones-set {Jones} will get the job.

Without delving too deeply into the theory of definite descriptions, let me consider one salient alternative to the Russellian view. Suppose we treat the definite description like a demonstrative, whose sole contribution to the sentence is to pick out a referent intended by the speaker, that referent being all that figures in the proposition expressed by the sentence (in some loose sense of "figuring in"). In this case we do not have a speaker but a believer, and what the believer undoubtedly intends to pick out by the definite description in formulating his belief is Jones. Thus, on this interpretation, believing that the person who will get the job has exactly ten coins in his pocket is equivalent to believing that Jones has exactly ten coins in his pocket. If *this* is what S believes, then it seems clear that we have no Gettier counterexample at all. S knows that Jones has exactly ten coins in his pocket.

A final Gettier counterexample is the following:

> *Sure-fire Match*. Sure-fire matches are guaranteed to light whenever struck, and S has tested dozens of them and always found them to light on the first strike. S is now holding a Sure-fire match that he believes will light when struck. Now, due to an extremely rare chemical impurity in this particular match, it cannot be lit by striking. Yet, S will strike it, and when he does it will light because of a random, quantum mechanical fluctuation occurring within the match at that very moment. Thus S truly believes that the match will light when struck, yet he does not know it (see Skyrms).

Believing that the match will light when struck is believing that a certain match-striking will be a match-lighting. Which match-striking? The actual one from among the class of all possible ones.[4] S believes that the actual match-striking will be a match-lighting. S's concept of the actual match-striking is of it as representative of the class of all possible match-strikings with respect to the property of being such that it will be a match-lighting (or

[4] Not *all* possible ones, of course; the match will not light underwater. The relevant class of possibilities in this case is limited to those in the "nearby modal vicinity," whatever that amounts to.

would be, if it is a merely possible match-striking). There is no information in S's individual concept of the actual match-striking with special bearing on the question of whether it will (or would) be a match-lighting. All information bearing on this question is equally contained in his concept of every possible match-striking as a possible match-striking. Thus, by the Inability Thesis, S does not believe the singular proposition normally expressed as "The actual match-striking will (or would) be a match-lighting." Rather, he believes the general proposition that all possible match-strikings will (or would) be a match-lighting. That is false, for there is a (nearby) possible match-striking—one not attended by the quantum mechanical fluctuation—that would not be a match-lighting.

Conclusion

I think it likely that the story I have told about representative concepts and their relation to beliefs can be applied to explain away all Gettier counterexamples in the same manner that I have used it to explain away the six presented in this paper. If so, we have good reason to believe the story, for it would figure into a highly theoretically satisfactory account of knowledge, one which analyzes knowledge as merely true belief. Let me conclude by emphasizing the limitations of my arguments. I do not take myself to have proven that knowledge is merely true belief (such a proof would require at least a story explaining away justification counterexamples, which story would probably be a lengthy project in its own right), or even that the story of representative concepts is unobjectionable as piece of psychological theory. My goal has been merely to demonstrate the potential viability of TB as a theory of knowledge, and to chart the very beginnings of a course for defending it against the standard counterexamples.

WORKS CITED

Gettier, Edmund. "Is Justified True Belief Knowledge?" *Analysis* 23 (1963): 121-123.

Goldman, Alvin I. "Discrimination and Perceptual Knowledge." *Journal of Philosophy* 73 (1976): 771-791.

Lehrer, Keith. "Knowledge, Truth, and Evidence." *Analysis* 25 (1965): 168-175.

Sartwell, Crispin. "Knowledge is Merely True Belief." *American Philosophical Quarterly* 28 (1991): 157-65.

---. "Why Knowledge is Merely True Belief." *Journal of Philosophy* 89 (1992): 167-180.

Russell, Bertrand. *Human Knowledge: Its Scope and Limits*. New York: Allen and Unwin, 1948.

Skyrms, Brian. "The Explication of '*X* knows that *p*.'" *Journal of Philosophy* 64 (1967): 373-389.