

## *Stick To What You Know*

JONATHAN SUTTON  
Southern Methodist University

### 1. Introduction

I will be arguing that a subject's belief that *p* is justified if and only if he knows that *p*: justification *is* knowledge. I will start by describing two broad classes of allegedly justified beliefs that do not constitute knowledge and which, hence, cannot be what they are often taken to be if my view is correct. It is far from clear what my view is until I say a lot more about the relevant concept or concepts of justification that concern me. The following section describes several concepts of justification that epistemologists have employed, and, in particular, identifies the two concepts of justification that I claim are coextensive with the concept of knowledge. One of those is the deontological conception of justification: I will be arguing that one ought not believe that *p* unless one knows that *p*. I imagine that the major opposition to my view will be that it is simply *obvious* that there are justified false beliefs, a feeling that I try to dispel in the lengthy section on concepts of justification before I finally get around to giving the main arguments in favor of my view. A view as unorthodox as mine demands more than a single argument: I offer four in the third section. Everyone allows that many people have many unjustified beliefs, and everyone has some unjustified beliefs, but such beliefs appear to be far more prevalent on my view than on more orthodox views. In the last section, I argue that unjustified beliefs, although widespread, are not quite as common as they might appear to be on my view.

### 2. Two Kinds of (Allegedly) Justified But Unknown Belief

#### 2.1 *The Unknown Unknown*

(It is a solecism to call a belief that does not constitute knowledge an 'unknown belief'. I hope that the convenience of the term outweighs its incorrectness.)

Suppose that I have entered the land of fake dollar bills. A counterfeiting operation has established itself so successfully in a nearby neighborhood that almost all the currency in circulation is fake, although I suspect nothing. If I believe that the fake \$10 bill that I was just handed in change is genuine, my belief is justified although false, orthodoxy maintains. Of course, I could be “lucky” (we will have more to say about such luck below) and end up with one of the few genuine bills in circulation. Because of my obliviousness to the counterfeiting operation and lack of awareness of how hard it is to come by a genuine bill in this neighborhood, I do not know that I have a \$10 bill in hand although I have a justified, true belief that I do, according to received wisdom—this is a Gettier case. Similar examples of (allegedly—often understood hereafter) justified false beliefs are easy to come by and occur often in actuality. Many or even most can be used to construct examples of justified true beliefs that actually occur at least sometimes.

Someone who has this variety of justified belief does not know that he does not know; I will hence call such beliefs ‘unknown unknown’. Upon minimal reflection, it is likely that I would believe that I knew that I had a genuine \$10 bill whether that belief is true or false. Whether I am a Gettier victim or not (whether my belief is true or false), if I came to know that I did not or had not known that I had a \$10 bill—for example, by being informed of the prevalence of counterfeit bills in the area—I would also lose justification for my belief and, if rational, I would give up the belief itself.

Discussions of Gettier cases often stress the good luck that the Gettier victim enjoys—it is a lucky accident that his belief is true. This good luck is subsequent to *bad* luck, however, which is at least as noteworthy, although much less often noted.<sup>1</sup> It was bad luck that I was in the land of fake dollar bills, or the land of fake barns, or that my colleague had just sold his Ford despite my having such good evidence that he owned such a car. The unknown unknown variety of justified belief, whether true or false, would have constituted knowledge but for unfortunate circumstances. (Of course, the false beliefs would not have been false in more fortunate circumstances.) The believer could have formed his belief in exactly the same way in different circumstances—indeed, the kind of circumstances that are far more common in actuality—and his belief would have constituted knowledge. I would know that I had a genuine \$10 bill if I formed a belief that I did in the same way that I do in the land of fakes in normal circumstances. If I formed a belief that there was a barn over there just by looking in that direction from a distance in normal circumstances in which fake barns are absent, I would know that there was a barn over there. And if my colleague actually did own a Ford and had generated my evidence that he did in the regular way, I would know that he owned a Ford on the basis of that evidence alone if I were to so believe.

It is very plausible that the good luck in Gettier cases is what leads us to deny that the victim knows despite having a true and (allegedly) justified belief.<sup>2</sup> It is good luck that his belief is true. It is just as plausible, I suggest, that it is bad luck that leads us to call the unknown unknown category of justified beliefs, including the Gettier cases, *justified*. It is because the victim would have known but for his bad luck that we keep his epistemic virtue intact by employing this label. This suggests that any genuine concept of justification at work in our so labeling this category of justified belief is parasitic on the concept of knowledge. We only understand what it is to be justified in the appropriate sense because we understand what it is to know, and can extend the notion of justification to non-knowers only because they are would-be knowers. We grasp the circumstances—ordinary rather than extraordinary—in which the justified would know. Justification in the relevant sense is perhaps a disjunctive concept—it is knowledge or would-be knowledge. In light of these considerations alone, the suggestion that justification is a more fundamental notion than knowledge, that justification is what is *really* important in epistemology, is dubious.<sup>3</sup> These comments are, of course, mere suspicions—suspicions that I hope will gain strength throughout this paper, and which we will return to in section 3.5.<sup>4</sup>

## 2.2 The Known Unknown

It is sometimes thought that if a proposition  $p$  is known to be sufficiently probable, then one is justified in believing that  $p$  even if one does not know that  $p$ . An example to which we will return in some detail below (section 4.2) in discussing the lottery paradox is the belief that one's ticket will not win the lottery. One does not know that one will not win the lottery, many will agree, but one is justified in believing that one will not win on the basis of knowing that it is extremely likely that one will not win. Someone who believes justifiably that he will not win and yet does not know that he will not also typically knows or can know upon minimal reflection that he does not know that he will not win.<sup>5</sup>

Hence, I call this second category of justified belief the known unknown beliefs. In section 4.2, I will examine and endorse an argument that beliefs that one will not win the lottery formed on probabilistic grounds are not justified; the argument's conclusion can and will be generalized to all such beliefs formed on merely probabilistic grounds.

There is another type of belief that is an example of known unknown allegedly justified belief, whose instances are formed by inference to the best explanation. Suppose that the total evidence one has supports one theory (which might be as small as a single proposition) over all rival theories, but that evidence is not sufficiently strong that one knows the theory to be true on the basis of that evidence. Many will say that that evidence might nevertheless be strong enough to justify belief in the winning theory whether it is true or false. (Although of course it *need* not be strong enough to justify

belief; the evidence might be too weak in either quantity or the degree to which it outweighs evidence for rival theories to do that.)

One who believes a proposition *p* on the basis of an inference to the best explanation that falls short of providing him with knowledge that *p* will often know that he does not know, or at least could do so upon minimal reflection. Just as one who forms a belief that *p* on probabilistic grounds in the relevant kind of case knows that *p* is probable, one who forms a belief that *p* through inference to the best explanation in the relevant kind of case knows *something*—namely that the evidence on balance supports *p* to a greater or lesser degree compared to such-and-such specified or unspecified rival theories.

I will have more to say about the known unknown class of allegedly justified beliefs in particular in section 4.3, arguing (of course) that they are not justified after all.

### 3. Five Concepts of Justification

#### 3.1 *Warrant*

It is tempting to begin with an apparently neutral definition of justification as what Plantinga calls *warrant*, leaving a more substantive characterization of justification to be developed after one has seen what fits Plantinga's definition (Plantinga, 1993a,b). Justification as warrant (Plantinga rejects the term 'justification' because of its deontological connotations) is whatever it is that makes the difference between mere true belief and knowledge—what has to be added to true belief to achieve knowledge. Knowledge is belief that is a) true and b) has property *X*, which is warrant/justification by definition.

The apparent neutrality of this definition is merely apparent, however. It assumes, firstly, that knowledge can be defined (and defined in an illuminating way, in more epistemically primitive terms) and, secondly, that such a definition will include a component that does not entail truth, that specifies a property that false beliefs can share with true beliefs. Both of these assumptions are questionable, and have been questioned. Linda Zagzebski and Trenton Merricks have argued that whatever differentiates mere true belief from knowledge entails truth (Zagzebski, 1996; Merricks, 1995), and Merricks notes that many definitions of knowledge, such as Nozick's (Nozick, 1981), do not include a separable warrant component that does not entail truth.<sup>6</sup> Timothy Williamson has argued that knowledge is not definable at all, and should be regarded as a conceptual primitive (Williamson, 2000). As he notes, the definability of knowledge does not follow from the fact that knowledge entails truth; being red entails being colored, but no one expects to define being red as being colored plus something else. We make an analogous assumption if we assume that there is such a thing as warrant in Plantinga's sense.

I am sympathetic to Williamson's position both on his grounds and for independent reasons, but I do not have the space to explore the view that knowledge is a primitive here. *If* Williamson, Zagzebski, and others are wrong and knowledge can be defined in terms of a notion like Plantinga's warrant, then clearly *my* notion of justification which I wish to identify with knowledge cannot be warrant. Necessarily, warrant is entailed by but does not entail knowledge. Equally, if my arguments succeed, then warrant is not even extensionally equivalent to either of the notions of justification that I do identify with knowledge. Some of those seeking to identify a notion of warrant such that warranted true belief is knowledge have wanted to identify warrant with the evaluative conception of justification outlined below. Since I will be arguing that evaluative justification is knowledge, at least some of my arguments need to be effective against at least these notions of warrant. One of my arguments will need to be modified to be so effective since it assumes that there are justified true beliefs that do not constitute knowledge; I will note the modification required. Some of my arguments simply do not apply to any notion of warrant.

The case against warrant that my arguments present is perhaps not as compelling as the case against other notions of justification, but, taken together, the arguments have enough persuasive force to rule out an evaluative notion of warrant. On the other hand, the notion of warrant is considerably less intuitive than a more orthodox notion of justification.<sup>7</sup> The unknown unknown beliefs, which provide intuitive support for the orthodox notions, do not provide intuitive support for the notion of warrant since those beliefs are *not* warranted. The unknown unknown true beliefs are *ex hypothesi* unwarranted since warrant solves the Gettier problem; *warranted* true beliefs constitute knowledge. Unknown unknown false beliefs are also, it seems, unwarranted since they have counterparts that are true. If the false beliefs were warranted, it is hard to see how their true counterparts would lack warrant and hence it is hard to see why warrant would not give rise to Gettier problems after all.

The known unknown beliefs are no better at providing intuitive support for the notion of warrant. True beliefs formed on probabilistic grounds or on the basis of inference to the best explanation that fall short of knowledge are unwarranted by definition; once again, it is very hard to see how their false counterparts could be warranted. In short, neither of the two classes of allegedly justified beliefs that are supposed to provide illustrations of justification without knowledge provide even apparent illustrations of warrant without knowledge. The claim that there *is* a notion of warrant, even if it is supposed to do double duty as a component in the definition of knowledge and as a notion of evaluative justification, is (unsurprisingly) no more intuitive than the claim that knowledge is definable as truth plus something else.

### 3.2 *The Deontological Conception*

A deontological conception of justification supposes justification to be tied to epistemic obligations—what one ought to believe, what one ought not believe, what one is permitted to believe, what one is not permitted to believe, what one is permitted not to believe, and so on.<sup>8</sup> Philosophers who have claimed that there are epistemic obligations have most commonly taken those obligations to be negative—there are certain beliefs that one should not have. For example, evidentialists claim that one should not have beliefs that are not supported by one's evidence (Feldman and Conee, 1985). Some philosophers have claimed that there are positive epistemic obligations, too: again, some evidentialists have claimed that one ought to believe what one's evidence does support (Feldman and Conee, 1985, again).<sup>9</sup>

I claim that beliefs justified in a deontological sense are those and only those that constitute knowledge. I claim that we human beings have a negative epistemic obligation: one ought not believe that  $p$  unless one knows that  $p$ , for any proposition  $p$ . Equivalently, one is permitted to believe that  $p$  only if one knows that  $p$ . I will not in general be concerned to argue that there are any positive epistemic obligations, and, indeed, I am inclined to think that the vast majority of beliefs that one ought to hold are such for non-epistemic reasons. Such positive doxastic obligations depend on what is important or interesting to oneself or others, among other considerations, and those notions of importance or interest are not epistemic notions. In section 4.3, I will in passing suggest one notable exception.

Many philosophers have assumed that to adopt a deontological notion of justification is also to adopt an *internalist* notion of justification (Sosa, 1999, for example)—indeed, that it is a deontological understanding of justification that provides a motivation or *the* motivation for internalism (Goldman, 1999; Plantinga, 1993b). In the next section, I will discuss to what extent my view of justification as knowledge is internalist, externalist, or uncommitted to either position. I will also discuss and endorse Hilary Kornblith's view that externalism and deontology are quite compatible, and apply his conclusions to my own view of justification.

Some philosophers, such as William Alston (Alston, 1985), do not believe in deontological justification, since they think that it entails a degree of voluntary control over what beliefs we form that we do not possess. Such arguments have been ably criticized by others in my view (Steup, 2001a; Kornblith, 2001; Feldman, 2001) and I will not consider them here, merely noting that I agree with these critics that we do not need any problematic level of voluntary control of our beliefs for there to be epistemic obligations. Indeed, I will not argue that there are any epistemic obligations (such a task seems as fruitless as arguing that there are moral obligations), simply that if there are any, then they have the direct relation to knowledge that I propose.

### 3.3 The Evaluative Conception

I identify evaluative justification with knowledge just as I did deontological justification. The remaining conceptions that I discuss below I consider primarily to distinguish them from these two, and to argue that they are much less central notions of justification, if indeed they deserve the name ‘justification’ at all.

Many have followed Alston in distinguishing the evaluative conception of justification from the deontological conception, of which he is suspicious as we noted above. A belief is justified in the evaluative sense if it is “a good thing from the epistemic point of view” (Alston, 1985, p. 329); there is no commitment to claiming that one ought to pursue beliefs that are a good thing from the epistemic point of view, or that one ought not pursue beliefs that are a bad thing from an epistemic point of view. At its most general, what is a good thing from an epistemic point of view for Alston at least involves aiming at maximizing true belief and minimizing false belief. He claims that this is “uncontroversial” (Alston, 1993, p. 535), and many philosophers have assumed without argument that truth maximization and falsity minimization are a primary or the primary epistemic goal (or goals). More must be said, Alston tells us, to identify a general epistemic goal that is *epistemic*. Truth, after all, is perhaps a supremely good thing from an epistemic point of view, but justification cannot be identified with truth since it is not *internal* enough, he says (Alston, 1993, 1985, again).

I claim that what is justified in the evaluative sense is knowledge—it is knowledge that is the supremely good thing from an epistemic point of view and, unlike truth perhaps, it is epistemic enough to *be* justification. Indeed, truth cannot be identified with justification for an entirely different reason from that which Alston cites—it has nothing to do with its lack of internal accessibility. The simple reason is that a true belief need not constitute knowledge, hence truth is not good enough for justification. Furthermore, it is not or should not be uncontroversial that our primary epistemic goal is a combination of truth maximization and falsity minimization.

The reason for this is that, as I will argue, the known unknown category of allegedly justified beliefs are not, in fact, justified. If a justified belief is one formed in serving the aim of maximizing true belief and minimizing false belief, the known unknown beliefs are apparently justified. If one knows (or could know upon minimal reflection) that it is highly probable that *p*—for example, that it is highly probable that one’s ticket will lose the lottery—then in forming the belief that one will lose, one serves the aim of maximizing truth while minimizing falsity, it appears, since it is by hypothesis highly probable that the belief one has formed is true. Similarly, if one knows that the best available evidence supports a theory *T* over its rivals to such a degree that *T* is (merely) *highly likely* to be true, then in forming a belief that *T* is true, although one’s evidence is not so good that one knows that to be the case, one is highly likely to have formed a true belief and so, it



appears, have served the aim of maximizing truth and minimizing falsity. This is not to say that there is *no* interpretation of truth maximization/falsity minimization consistent with saying both that this goal somehow determines or is closely connected to evaluative justification and that the known unknown beliefs are not really justified. It is to say that it should not be uncontroversial that there is such an interpretation—it should not be assumed that there is one. I will offer an alternative account of what is a good thing from the epistemic point of view in what follows.

Another assumption is almost as common as the assumption that truth maximization/falsity minimization is a primary epistemic goal—the assumption that a central fact about belief is that it *aims* at truth (e.g., Williams (1973, p. 148, cited in Ginet (2001))). Known unknown beliefs again suggest that this is not so.<sup>10</sup> If belief aims at truth, then the belief that one will lose the lottery or that a theory such as *T* is true will, in almost all cases, succeed in fulfilling that central aim, and so should be impeccably formed, i.e., justified. If, as I will argue, the known unknown beliefs are *not* justified, and not justified because they do not constitute knowledge, we should rather say that belief aims at knowledge.<sup>11</sup>

There are many truths that it is not in one's interests to believe *or know* because they are unimportant, uninteresting, tasteless, upsetting, or dangerous. One needs to square this fact with the claim that truth maximization is a primary epistemic goal if one wishes to support that claim. One is likely to do so by claiming that it is in one's *epistemic* interests to believe such truths, but not in one's non-epistemic interests which can easily outweigh one's epistemic interests. But there is no need to cast epistemic goals in terms of *maximizing* anything once we have abandoned naked truth as the center of our epistemic life. We can simply talk of the aim or goal of individual beliefs and of belief in general. Alston (Alston, 1993) says that we cannot define justification as what serves our most general epistemic goals since that is truth, which is clearly not what justification amounts to (since it is not "epistemic enough"). But we can so define justification if belief aims at knowledge rather than truth. There is no temptation to bring truth into our account of justification in a suitably *epistemic* way by refining the notion of truth maximization since it is entailed by knowledge, a paradigm of the epistemic.

Talk of the *aim* of belief is more than a little obscure, as is talk of distinctively *epistemic* goals itself. I would like to clarify such talk by saying that the *function* of belief is to be knowledge; that is what belief is *for*, in terms of whatever notion of design or purpose one wants to apply to human beings, their states, and faculties. Belief can serve many functions, but *the* function—the proper function, if you like—is to be knowledge. And the proper function of the faculties that produce beliefs is to produce beliefs that constitute knowledge. If this is right, defining justification or warrant in terms of the proper function of those faculties or of the beliefs themselves is just to define justification as knowledge rather than a more primitive *component* of knowledge (namely, warrant), contra Plantinga (Plantinga, 1993a).



How do I propose to argue that justified belief in the evaluative *and* deontological senses is knowledge? A number of my arguments establish a conclusion stateable using either conception of justification (the assertion and lottery arguments in particular). All of my arguments indirectly support a dual conclusion, however, since the identification of one of the two forms of justification with knowledge gives rise to an argument that the other form must also be identified with knowledge. Firstly, let us see that if deontological justification is knowledge, then so is evaluative justification.

I shall assume that justification in any important evaluative sense does not require *more* than knowledge—too few of our beliefs would be justified on a more stringent conception for the notion of evaluative justification to play the important epistemic role that it is supposed to play.<sup>12</sup> Suppose that it requires less than knowledge. Suppose that my belief that *p* is justified in the evaluative sense, that it satisfies a primary or the primary epistemic goal, but that I do not know that *p*. Suppose further that deontological justification is knowledge, and so my belief is not justified in the deontological sense. Then, I ought not believe that *p* although I am satisfying primary epistemic goals in so doing. That seems completely mysterious.<sup>13</sup> If there are epistemic obligations at all, it is conceivable (although false, I suggest below) that fulfilling primary epistemic goals requires going beyond one's epistemic obligations, but it is inconceivable that one might simultaneously fulfill those goals and violate those obligations. If there are epistemic obligations at all, one is surely epistemically *permitted* to fulfill primary epistemic goals. Indeed, it is quite plausible that one is obligated to *pursue* those goals in some sense even if one is not obligated to fulfill them. So, I contend, if epistemic obligations are as stringent as I claim that they are, the deontological and evaluative conceptions of justification coincide.

Now I will argue that if evaluative justification is knowledge, then deontological justification is, too. This argument relies on the assumption that there *are* epistemic obligations, and something must first be said about what such obligations require.

Typically, as noted above, deontological conceptions have been associated with internalist theories of justification. Is my view externalist or internalist? Strictly speaking, it is neutral. Since justification is identified with knowledge, it is as externalist or internalist as knowledge itself, on which I do not *need* to take a position for most of what I say in this paper to succeed. However, very few philosophers nowadays hold an internalist view of knowledge. Almost all would agree that there are possible (and likely actual) pairs of thinkers (interworld or intraworld) who form a belief with a given content *p* and who are in exactly similar relevant mental states—and have *access* to all the same relevant mental states—one of whom knows that *p*, the other of whom does not.<sup>14</sup> In section 2.1, I suggested that all the unknown unknown allegedly justified beliefs can provide examples. Someone with such a belief would have known if entirely external circumstances had been different.<sup>15</sup> Consequently, one does not have *infallible*

access to whether or not one knows.<sup>16</sup> But only a militant Cartesian would claim otherwise, and there are few of those left.

On the other hand, my view avoids some of the classic problems for the most prominent externalist theory of justification, reliabilism. My view does not even appear to entail that a clairvoyant who doubts that he is capable of reliably forming the beliefs that he does in fact reliably form forms justified beliefs, since it is very plausible that such a clairvoyant's beliefs do not constitute knowledge (Bonjour, 1985). Indeed, it is obviously consistent with my view that no clairvoyant or chicken-sexer, however confident in his beliefs, forms justified beliefs since it is consistent with my view that these characters' beliefs do not constitute knowledge.

Although it is common to associate deontological conceptions of justification with internalism about justification, and many externalists about justification (of which I am an exotic example, given the above) disavow any epistemic obligations, Hilary Kornblith is a notable exception (Kornblith, 2001). Kornblith argues that we should understand justification as an *ideal* state, one determined by the nature of the human mind and its capacities, but not any *individual's* mind or its capacities. Some—perhaps all—humans will be more or less incapable of living up to that ideal in some of the situations in which they form (or fail to form) or renounce (or fail to renounce) beliefs.<sup>17</sup> He says:

An appropriate human ideal must in some ways be responsive to human capacities. Ideals are meant to play some role in guiding action, and an ideal that took no account of human limitations would thereby lose its capacity to play a constructive action-guiding role. At the same time, our ideals cannot be so closely tied to what particular individuals are capable of that we fail to recognize that some individuals at some times are incapable of performing in ideal ways. There is a large middle ground here, and it is here that reasonable ideals are to be found. (Kornblith, 2001, p. 238)

Evaluative justification is a matter of satisfying primary epistemic goals—or, as we might just as well say, ideals. Provided that we understand those ideals as *humanly* possible, at least in principle, and we adopt Kornblith's idealized understanding of deontological justification, we have a way to argue from claims about evaluative justification to counterpart claims concerning deontological justification.

On my view, we ought to live up to the ideal—knowledge—that is the aim of belief, however difficult (or even “impossible” *in practice*) that might be in particular circumstances, and however much constant struggle and vigilance is accompanied by inevitable lapses. All of which brings us to the question of the relationship between blame and obligation, and to our next conception of justification. I will in section 6 return to Kornblith's claim that justification plays an action-guiding role.

### 3.4 *Justification As Blamelessness*

Externalist theories of justification entail that many actual and possible believers have beliefs for which they cannot be blamed which are nevertheless unjustified. Goldman (Goldman, 1988) gives the example of believers who grow up in a “scientifically benighted” society who form some of their beliefs—beliefs, moreover, that are crucial to whether they succeed or fail to achieve goals important to them—by worthless manipulation of zodiacal signs. Similar examples are easy to construct. There is almost no end to the crazy conclusions at which one might routinely arrive whose craziness it is practically impossible for one to recognize because it is inconceivable for people in one’s social circumstances to believe anything different. Another problem that externalists worry about is what Sosa (Sosa, 1991) calls ‘the new evil-demon problem’. On many versions of reliabilism, a thinker globally deceived by an evil demon or who is a brain in a vat will have very few beliefs that are justified since they are not arrived at by the relevant kind of reliable process.

Goldman is happy to define a notion of a *weakly* justified belief that is more or less coextensive with blameless yet ill-formed belief (strong and weak justification are contraries) and which the zodiacal believers (and brains in vats) live up to, reserving ‘strong justification’ for justification in the epistemically important sense that concerns him. There are indefinitely many properties that some beliefs have and others lack that render the beliefs that have them good in some way connected more or less strongly with arriving at the truth. Goldman’s strong and weak justification are two points on this continuum (even the zodiacal believers do not believe that they are forming beliefs via a process with an excellent chance of arriving at false belief, a “virtue” captured by Goldman’s definition of weak justification). Hartry Field (Field, 1998) has noted that attempts to pick a point on this continuum and regard it as genuine justification give the impression that one is searching for some “justificatory fluid” that is sprayed on some beliefs and not others. He embraces the view that none of these competing notions are objectively any better than the others, a view he calls ‘epistemological non-factualism’.<sup>18</sup> (It is the position of this paper that an emphasis on truth-maximization is largely responsible for Field’s scepticism. The evaluative and deontological conceptions of justification are the ones that matter, and they do specify a unique justificatory fluid—knowledge.)

I am inclined, then, to regard Goldman’s position as overly concessive. Its concessiveness is largely terminological, but it is the kind of terminological choice that leads one to carve out a multitude of notions of “justification” in logical space, and ultimately to Fieldian scepticism. Better to say that blamelessness is one thing and justification another, and that to call a belief unjustified is not to call it blameworthy. The new evil-demon problem is generated by a failure to recognize the distinction. For sure, if there are epistemic obligations at all, one *ought* to form only epistemically blameless

beliefs, but this obligation should arise as a consequence of a more stringent obligation (as it certainly does on my view). Blamelessness can hardly be taken to be a primary epistemic goal—any purported epistemic goal that the zodiacal believers live up to qua zodiacal believers is much too modest to be central to our epistemic lives.

An externalist account of justification that endorses the existence of epistemic obligations faces the new evil demon/benighted society “problems” to an even greater degree. It is conceivable that there is *no* connection between evaluative justification and blame, but there is surely *some* connection between deontological justification and blame. If you do not do what you ought to do, then, in at least some situations, you can be held to account on that basis.

Surely no externalist about justification faces this kind of problem as acutely as I do. I am not casting apparent aspersions solely on brains in vats or readers of runes, but on every believer at least some of the time. Here is perhaps the worst case. We can have justified beliefs about the future, which on my view amounts to the fact that we can know many things about the future. You believe that you will meet me at the airport tomorrow, and act on this belief in typical ways—you tell me that you will meet me there, and so on. But you drop dead this evening. I will not hold you to your obligation to *act*, to pick me up at the airport. But you did violate your epistemic obligations, on my view. You did not know that you would pick me up at the airport (although you would have known in more fortunate circumstances), and so, according to my view, you should not have believed that you would so act. How callous!

Such examples need not involve the future, of course. You believe that I live in a city. But *I* dropped dead five minutes ago, so I do not live anywhere, and your belief is false and so does not constitute knowledge. You ought not believe that I live in a city, although you once knew that I did (but ten minutes ago) and you could not be *expected* to be aware of the change in circumstances. Indeed, if I dropped dead a sufficiently small time ago, it will be *physically impossible* for you to realize that you must change your beliefs so as to live up to your epistemic obligations. That I regard standard examples of Gettier victims (and their unlucky counterparts with false beliefs) as in violation of their epistemic obligations seems almost trivial by comparison.

(It is rather strange for *any* externalist to regard the unknown unknown beliefs as justified (we have already seen that proponents of warrant cannot so regard them, largely by design). What determines evaluative justification for an externalist is adherence to a set of epistemic standards that are external to the believer’s mind. An unknown unknown belief is justified in part because of *ignorance*. The naïve possessor of a genuine or fake dollar bill in the land of fake dollar bills is allegedly justified in believing his bill genuine only because he is ignorant of the local counterfeiting operation. If

he became aware of that operation, he would no longer be justified in believing that his bill was genuine. It is a puzzling *external* epistemic goal that one achieves through ignorance but which one fails to meet by subsequently acquiring knowledge. And this is not a special case—*all* of the unknown unknown allegedly justified beliefs have this characteristic. An internalist—for example, a coherentist—has no problem here, since the internalist's epistemic standards are entirely determined by one's internal states, which obviously change when ignorance is removed, and in ways that clearly can and often do deprive, for instance, a formerly coherent set of beliefs of their coherence.<sup>19)</sup>

With the blamelessness/justification distinction in hand, a deontological externalist has a ready response to the new evil-demon problem. A deontological externalist should be as much of a fallibilist about epistemic obligations as he is about justification itself. Brains in vats and zodiacal believers are not in a position to know that their beliefs are unjustified—nor are they in a position to know what they should (or, more relevantly, should not) believe. Brains in vats and zodiacal believers cannot—at least some of the time—live up to epistemic ideals or recognize what those ideals demand (although they might *think* that they can) and we cannot blame them for that. This does not remove their obligations. Often we have to teach people, particularly children, what their obligations (epistemic and otherwise) are, and often they are not capable of recognizing or living up to those obligations until we do. But we do not *create* those obligations by our instruction; we help them by making them aware of obligations they have that are already in effect. The obligation/blameworthiness distinction is needed to make sense of instructing people in what they ought to do—or believe.

My kind of deontological externalism simply extends this story about brains in vats and zodiacal believers to all of us, at least some of the time. Just as a conflation of the distinction between blamelessness and justification can explain why some feel that a brain in a vat has many justified beliefs, that same conflation can explain the judgment that the ignorant in the land of fake dollar bills justifiably believe that they have genuine dollar bills.<sup>20</sup> We are all such that we sometimes are mistaken about our epistemic obligations and in such a way that it is in practice impossible for us to recognize that this is so. In practice, but not in principle—as Kornblith says, our epistemic obligations are ideals, but they are constrained by human capacities. Someone in the land of fake dollar bills is in principle capable of recognizing that he does not know that his bill is genuine (he just needs to check out the criminality of his neighborhood a little more thoroughly) just as the brain in a vat is in principle capable of living up to his epistemic obligations (he just needs to get out of his vat and get a life). You could have lived up to your obligations with regard to your belief that you would pick me up at the airport by *not dying* just as a zodiacal believer could, in ideal conditions, come to realize that his astrological upbringing was

worthless. The constraints placed upon our epistemic obligations by human capacities are very loose in the cases that are problematic for any deontological externalist—but the airport and fake bill cases are not *more* of a problem for me than the brain in a vat case is for any deontological externalist. Some possible and actual situations prevent us from realizing what our obligations are, but most of us are not in those situations most of the time, and it is humanly possible to get out of these situations or not get into them in the first place. I do not, after all, require that everyone should believe all the logical consequences of their beliefs as certain versions of evidentialism notoriously do (Feldman and Conee, 1985).

A fallibilist about knowledge should say that even if we cannot infallibly know whether or not we know, we *often* can tell whether we know. A fallibilist about epistemic obligations should say the same. On my view, this is easy. Since we can often know that we know or know that we do not know, we can often know what we ought not believe—that which we do not know. The connection between epistemic obligation and what humans are capable of *in practice* is a loose fit, but the two go together for most of us most of the time. Which brings us to the connection between *praiseworthiness* and epistemic obligation. Once again, the connection is loose: we do not praise people for failing to form any of a number of unjustified beliefs that they could form. One can, however, say this: if someone violates his obligations, he is not generally praiseworthy. Forming known unknown allegedly justified beliefs is a violation of epistemic obligations on my view. Consequently, believing that one has not won the lottery or that one will not die in a wreck on the highway or some scientific (or philosophical) theory that is best supported by one's evidence but not to such an extent that one knows it to be true is not generally praiseworthy.

Undoubtedly, such unjustified beliefs are commonplace, as are their unknown unknown counterparts. This is not an objection to my view. Everyone thinks that unjustified beliefs are common, and everyone thinks that many people violate their obligations (of whatever kind) on a regular basis. As in any branch of philosophy that attempts to delineate our obligations (ethics being a central example), we are only loosely constrained by what others (philosophers and non-philosophers) *take* to be their obligations, and we are only loosely constrained by the observed frequency with which people live up to what we conclude their obligations are. It is quite conceivable that regular violations are the norm in some epistemic situations, just as in some moral situations. Moreover, in epistemic situations as in moral situations, we should expect that the violations of some will encourage the violations of others, leading to localized epistemic scandals to which the protagonists are oblivious.

### 3.5 *Justification as Reasonableness or The Bit Where You Take It Back, Part I*

It has been noted that 'justified' is rarely, if ever, applied to beliefs in everyday speech, the term being more commonly applied to actions, and



that ‘reasonable’ is a term that is used of beliefs in much the way that philosophers use the term ‘justified’. Surely, one might object at this point, people are talking about *something* when they talk of reasonable yet false (and therefore unknown) beliefs.

I agree. This is no reason to think that reasonableness can be identified with justification in the philosophers’ sense, however. For justification is supposed to be primarily a property of beliefs, and reasonableness can be understood as only a property of beliefs in a derivative sense. Reasonableness is, I suggest, primarily a property of *persons*—a belief is reasonable in the circumstances in which it is held if a reasonable person would or could hold it in those circumstances (shades of “responsibilist” virtue epistemology (Zagzebski, 1996)). The notion of a reasonable *person* is understood in terms of knowledge; he is one whose belief-forming faculties and habits (e.g., inferential habits) are such as to deliver knowledge when conditions are right. Of course, reasonable people are sometimes, through no fault of their own, in the wrong conditions, and then they form unjustified beliefs—unjustified on my view since they do not constitute knowledge. Those beliefs are nevertheless reasonable by definition. One who aims at knowledge and who is successful for the most part thereby achieves reasonableness—reasonableness is not an independent epistemic goal (it is not evaluative justification) and nor is there an obligation to be reasonable distinct from the obligation to believe only if one knows (it is not deontological justification). Not everyone who believes blamelessly is reasonable; remember the zodiacal believer—reasonableness is not blamelessness. (Neither, obviously, can it be used to define knowledge—it is not warrant.)

This characterization of reasonableness is very much of a piece with my suggestion in section 2.1 that our understanding of the notion of justification allegedly present in the unknown unknown beliefs is parasitic on our understanding of knowledge—it is would-be knowledge, so to speak. Known unknown beliefs are simply not reasonable even in this sense since they would not constitute knowledge even if conditions were right. (Although those who form beliefs in theories that they do not know to be true might be blameless since they are entrenched members of an unreasonable community, much like Goldman’s zodiacal believers.)

(Are the beliefs of recently envatted brains reasonable in the current sense? Are the beliefs of native brains-in-vats reasonable? I imagine that the notion of reasonableness, even on the assumption that it is one employed by “the folk,” is far too vague for these questions to have determinate answers. The notion of a reasonable person does not determinately express a single notion definable in terms of knowledge. It exhibits the “nonfactualism” that Field attaches to the more central evaluative conception of justification.)

The word ‘justification’ is perhaps connected with belief in common speech not as a *property* of a belief, but in the notion of someone’s justification *for* a belief. Such justifications can be adequate or inadequate;



I contend that an adequate justification for a belief is one such that, if the belief were formed on its basis, it would constitute knowledge. A belief that has an adequate justification for an individual, then, is once again to be identified with knowledge. An inadequate justification is one that fails in its aim—it does not render that for which it is a justification known.

#### 4. The Arguments

The four arguments in this section stand alone; each is supposed to establish its conclusion without help from the others. Nevertheless, they also work together in demonstrating a number of diverse phenomena that are best explained by the hypothesis that justification is knowledge.

##### 4.1 *The Assertion Argument*

Philosophers including Peter Unger (Unger, 1975), Michael Slote (Slote, 1979), Keith DeRose (DeRose, 1996), and Timothy Williamson (Williamson, 2000) have argued on broadly similar grounds that one is not warranted in making an assertion that  $p$  unless one knows that  $p$ .<sup>21</sup> The notion of warrant involved is explicitly deontological for Williamson (and for Unger) who says that “one must: assert  $p$  only if one knows  $p$ ” (Williamson, 2000, p. 243), a norm—the knowledge rule—the following of which he takes to be constitutive of assertion. In this section I will review Williamson’s main arguments for the claim that warranted assertion requires knowledge, which are the most exhaustive of the four authors. I will then argue that the entailment from warranted assertibility to knowledge is inexplicable unless there is a similar entailment from justified belief to knowledge. One must: believe  $p$  only if one knows  $p$ , a norm that is as constitutive of belief as Williamson’s knowledge rule is of assertion.

I shall divide Williamson’s arguments for his knowledge rule into three, further discussion and defense of which can, of course, be found in Williamson’s text. The first states that the knowledge rule is required to explain why quasi-Moorean assertions of the form ‘ $p$  and I don’t know that  $p$ ’ are deeply unacceptable. Sentences of this form could clearly be *true*; the knowledge rule explains why they implicate something that clearly contradicts what is explicitly asserted. An assertion is presumed to obey the norms governing assertion. We assume, then, that the asserter knows the asserted conjunction, which entails that he knows the first conjunct, i. e., that he knows that  $p$ , which is explicitly denied in the second conjunct.

The second argument is that lottery cases seem to show that evidential standards falling short of knowledge are not sufficient to warrant assertion. If one buys a lottery ticket, no matter how probable it is that one’s ticket will lose, and no matter how aware one is of that probability, one is not warranted in asserting that one will not win. Almost everyone acknowledges that one does not *know* that one’s ticket will lose; the knowledge rule, then,

explains the unacceptability of the assertion that one will lose. The third argument is simply the appropriateness (considerations of politeness and tact aside) of challenges and rebukes to assertions such as 'How do you know?' and 'You don't know that!'

Now, let us suppose that Andy has a justified true belief that  $p$  that does not amount to knowledge that  $p$ . (This is impossible on some views of justification, namely those that take justification to be Plantinga's warrant—whatever property that, combined with truth, renders a belief knowledge. I will extend the argument to such views below.) We might take ' $p$ ' to be a member of the unknown unknown beliefs, e. g., that Andy is giving Bob a dollar bill, believed in the land of fake dollar bills, or a member of the known unknown beliefs, e. g., that Andy's ticket will not win the lottery. By the knowledge rule, it is impermissible for Andy to assert that  $p$  to Bob, but let us suppose that Andy makes this unwarranted assertion, and that Bob believes what he is told. Let us moreover suppose that both Andy and Bob's beliefs that  $p$  are justified by whatever alleged standards of justification one cares to select (externalist or internalist, deontological or evaluative) falling short of knowledge consistent with the case.<sup>22</sup> For example, we might suppose that Andy formed his belief by a process that fairly reliably produces truths *if* a true belief could arise by such a process that was justified but did not constitute knowledge (reliabilists might differ on this point). We can suppose that Bob's reliance on Andy's testimony is similarly reliable. Indeed, we might suppose that Andy *only* utters truths in actual and a wide range of counterfactual circumstances, so that Bob's belief perhaps has an even higher degree of justification than Andy's, although he cannot know that  $p$  on the basis of Andy's testimony unless Andy knows that  $p$ .<sup>23</sup> We can also suppose that Andy's belief is impeccably justified by whatever internalist standards one likes consistent with his not knowing that  $p$ , and that Bob's reliance on Andy's testimony is similarly free of internalist blemish.

More generally, if there *is* a gap between justified true belief and knowledge, we can suppose that that gap is exemplified in the case of Andy and Bob; the details will vary wildly, of course, depending on the particular theory of justification we select. Andy has a justified true belief that  $p$  which does not constitute knowledge and he asserts that  $p$  to Bob who has, we can suppose, the very best reasons for thinking—indeed, he knows—that Andy is expressing what is for him a justified belief.<sup>24</sup> Bob, then, has acquired a true belief that  $p$  that is justified by any number of standards that one likes consistent with it falling short of knowledge. Bob's belief violates none of his epistemic duties, and it lives up to primary epistemic goals; it is, we can further suppose, a *good thing* in numerous ways for Bob to believe that  $p$ . He will be healthier, wealthier, and wiser as a result. And yet, the knowledge rule tells us, Andy should not have asserted that  $p$ .

This is exceptionally puzzling. One of the main goals of making assertions, if not *the* main goal, is to transmit beliefs from one thinker to

another.<sup>25</sup> If the beliefs so transmitted meet the primary standards governing good belief for both speaker and hearer—that is, they are justified in an evaluative sense—and meet standards of permissible belief (as noted previously, it is hard to imagine the former standards being met without the latter), it would be mysterious if the assertions transmitting the beliefs failed to meet the standards governing good assertion. On the contrary, the assertions in question have to meet the standards governing good assertion impeccably since they transmit impeccable beliefs. It is not, however, the knowledge rule that is at fault; the arguments of Williamson and others for that rule are good ones. It is our initial supposition that was at fault. There are no justified true beliefs falling short of knowledge, so Andy cannot have one. Assertions that transmit impeccable beliefs are impeccable assertions according to the knowledge rule since all impeccable beliefs constitute knowledge.

What, then, if there *are* no justified true beliefs that fail to constitute knowledge; what if the only workable notion of justification is a notion of warrant that in combination with truth suffices for knowledge? I suggest that my argument can be modified to retain its effectiveness, although it is perhaps a little diminished. (The modified argument also works as a supplementary argument that justification is knowledge for those who are not proponents of warrant.) We now must suppose that Andy transmits a warranted belief that falls short of knowledge to Bob, who satisfies all the other conditions specified above (for example, he knows that Andy is expressing a belief that is warranted for him). That is to say, Andy transmits a warranted *false* belief to Bob; Bob ends up with a warranted false belief in the same proposition in which Andy has a warranted false belief. Once again, we ask the question: how could an assertion that transmits a warranted belief from Andy to Bob be an unwarranted assertion when one of the main goals of assertion is precisely to transmit belief?<sup>26</sup> To which question it seems that there is an obvious answer: Andy told Bob something false, although his belief in it was warranted. Of course there was something wrong with his assertion.

The answer is a little too glib, however. Andy and Bob's beliefs are also false, there is also something wrong with them, but they are warranted. Why should the standards for warranted assertion be so much stricter than the standards for warranted belief when one of the main goals of assertion is to transmit belief? Once again, I suggest that they are not, and that is because a counterpart knowledge rule applies equally to belief.

One might respond that my characterization of the goal of assertion is simply mistaken. It is not one of the goals of assertion to transmit belief; in light of the knowledge rule, it is clearly to transmit knowledge. Indeed, one might go further and claim that Bob cannot himself acquire a warranted belief from Andy unless Andy expresses knowledge. The warrant of no warranted false belief can be transmitted to another via testimony. If that

is so, then the belief that Bob acquires will be as defective as Andy's assertion, and there is harmony between the standards of belief and assertion.

These defenses do not remove our initial puzzle, however. If a belief can be warranted despite being false, it is bizarre that an assertion conveying such a belief is unwarranted. Why do the main goals of assertion fail to encompass the conveying of such beliefs? Why is an assertion not doing its job when it conveys such a belief? It would add to our bafflement if Bob could not acquire a warranted belief from Andy's expression of a warranted false belief. Why would assertion not act as a conduit for warranted belief as well as knowledge proper? The best explanation of such phenomena would be that warranted belief just is knowledge: that is why Bob only acquires a warranted belief if Andy expresses knowledge, for only in those conditions is Andy's belief itself warranted (there being no such thing as warrant in Plantinga's sense of that term, if it is supposed to be evaluative justification as well as a component of knowledge).

It will also do no good to claim that what is wrong with Andy's assertion, whether one considers the original case or the modified justified false belief case, is that Andy potentially misleads Bob into thinking that he *does* know since that is implicated by his having made the assertion in the light of the knowledge rule. It is very strange that there should *be* any such implication if there are justified beliefs falling short of knowledge since it is hard to see how an assertion that transmitted such beliefs would be defective. One cannot appeal to the knowledge rule itself in trying to explain away the odd combination of justified beliefs that do not constitute knowledge and the knowledge rule; that is to ignore the oddity of the combination, not explain it away.

I will close this section with a new line of thought that bolsters my case that the knowledge rule is only explicable if justification is knowledge. The knowledge rule is arguably subject to exceptions, since in some cases it is arguable that speakers regularly assert falsehoods to communicate truths. It is sometimes suggested that so-called restricted quantification is one such example; when I assert (falsely, it is suggested) that everyone has been told of the meeting I might communicate the truth that everyone in my workplace has been told of the meeting.<sup>27</sup> If such a phenomenon occurs, it is surely pervasive, and we must suppose that such assertions are warranted. What is notable is that the falsehoods asserted in such situations would neither express the literal content of the speaker's belief, which would be the expanded counterpart truth, nor would they have any tendency to induce a belief in the falsehoods in the audience, who would also form a belief in the expanded counterpart truth. That is, these assertions do not induce unjustified beliefs in their audience (nor do they express unjustified beliefs of the speaker). This suggests that the knowledge rule governing assertions is perhaps an imperfect expression of the relation between belief and assertion.

One should assert only what one justifiably believes when there is otherwise a risk of inducing unjustified beliefs in one's audience. If justification is knowledge, this generalization explains both the knowledge rule that governs most assertions and the possible exceptions described above. The counterpart to the knowledge rule governing belief is what is fundamental.<sup>28</sup>

#### 4.2 *The Lottery Argument*

This argument builds on the work of Dana Nelkin (Nelkin, 2000), who presents two versions of the lottery paradox, one concerning knowledge and the other justification, and argues that they should receive a solution that locates the flaw in each paradoxical argument in the same place (in fact, counterpart premises).<sup>29</sup> I will review the two paradoxes as Nelkin presents them, then review her reasons for faulting the premises that she does. Nelkin's uniform solution to the paradoxes involves denying that the belief that one's ticket will not win is justified. I will argue that Nelkin's explanation of why it is unjustified is incomplete at best. The best explanation, I will argue, is that justification is knowledge.

##### **The Knowledge Paradox**

1. Jim knows that his ticket t1 will lose.
2. If Jim knows that his ticket t1 will lose, then he knows that t2 will lose, he knows that t3 will lose . . . and he knows that t1,000,000 will lose.

So,

3. Jim knows that t1 will lose . . . and Jim knows that t1,000,000 will lose. (1, 2)
4. Jim knows that either t1 will not lose or t2 will not lose . . . or t1,000,000 will not lose.
5. Propositions of the following form comprise an inconsistent set: (a) p1 . . . (n) pn, (n + 1) not p1 or . . . not pn.

So,

6. Jim knows propositions that form an inconsistent set. (3, 4, 5)
7. It is not possible to know propositions that form an inconsistent set.

So,

8. (1), (2), (4), (5), or (7) is false.

(Nelkin, 2000, p. 374, obvious typographical error corrected)

##### **The Justification Paradox**

1. Jim could justifiably believe that his ticket t1 will lose.<sup>30</sup>

2. If Jim could justifiably believe that his ticket  $t_1$  will lose, then he could justifiably believe that  $t_2$  will lose, he could justifiably believe that  $t_3$  will lose . . . and he could justifiably believe that  $t_{1,000,000}$  will lose.

So,

3. Jim could justifiably believe that  $t_1$  will lose . . . and Jim could justifiably believe that  $t_{1,000,000}$  will lose. (1\*, 2\*)
4. Jim could justifiably believe that either  $t_1$  will not lose or  $t_2$  will not lose . . . or  $t_{1,000,000}$  will not lose.
5. Propositions of the following form comprise an inconsistent set: (a)  $p_1$  . . . (n)  $p_n$ , (n + 1) not  $p_1$  or . . . not  $p_n$ .
6. Jim recognizes that the following propositions form an inconsistent set: (i)  $t_1$  will lose . . . (n)  $t_{1,000,000}$  will lose, either  $t_1$  will not lose . . . or  $t_{1,000,000}$  will not lose.

So,

7. Jim could justifiably believe inconsistent things that he recognizes are inconsistent. (3\*, 4\*, 5\*, 6\*)
8. One cannot justifiably believe things that one recognizes are inconsistent.

So,

9. (1\*), (2\*), (4\*), (5\*), (6\*) or (8\*) is false.

(Nelkin, 2000, p. 375, obvious typographical error corrected, and claims about what “it is rational for Jim to believe” replaced by claims about what he could justifiably believe)

Our conclusions below will apply whether the premises of the justification paradox are understood deontologically or evaluatively. (Premise (1\*), for example, can be paraphrased as “it is permissible for Jim to believe that his ticket  $t_1$  will lose” and premise (9\*) as “One ought not believe things that one recognizes are inconsistent,” or the notion of justification employed throughout can be understood in terms of the satisfaction of epistemic goals.)

Nelkin notes that almost everyone thinks that premise (1) is at fault in the knowledge version of the paradox; one simply does not know that one’s ticket will lose however aware one is of the very small chance that it has of winning. She notes that the justification version of the paradox rarely inspires such a reaction, prompting unintuitive positions that block the transferral of justification from premises to conclusions of apparent paradigms of good inference. For example, Foley (Foley, 1979) denies (8\*) by denying the conjunction rule that belief in a conjunction is justified if belief in each conjunct is. However, the difference between the solutions that have been offered to the two versions of the paradox is in part because both versions are rarely discussed together, yet it is a clear desideratum of a

solution to one version that it can be applied to the other. However uncomfortable it is to deny the conjunction rule for justification, it is far harder to swallow the claim that it does not hold for knowledge, which a parallel solution to the knowledge paradox would require. Denying premise (1) seems to be by far the best strategy with respect to the knowledge paradox, and Nelkin advocates denying (1\*) on that basis. Indeed, it is not remotely unintuitive to deny (1); the knowledge version of the paradox is not really a paradox at all, but simply an unsound argument. This does not undercut the force of Nelkin's point that the two unsound arguments ought to receive the same diagnosis, even if only one is a genuine paradox.

Of course, we need an explanation of why premise (1\*) of the justification paradox is incorrect. I will review Nelkin's explanation below. My explanation is simple. Premise (1\*) is false because it *is* premise (1); in a fundamental sense, the justification paradox just is the knowledge "paradox," and the solution to the paradox is to see that it is just the "paradox," which is nothing more than an unproblematically unsound argument. Justification is knowledge. Hence, to say that Jim could justifiably believe that his ticket *t1* will lose is to say that he could know that his ticket will lose—that he could know given just the grounds for belief that he already has, that if he were to believe, then he would know.<sup>31</sup> Premise (1\*) is false because premise (1) is false, since they express more or less the same proposition (ignoring the modal phrasing of (1\*)). The identification of justification and knowledge receives support from the fact that it enables a parallel diagnosis of the knowledge paradox and the justification paradox.

Such support is weakened if there is some other viable explanation of why (1\*) is false. Nelkin aims to provide just such a diagnosis.<sup>32</sup> Firstly, she notes that many externalist theories of knowledge require that there be a "causal or explanatory" connection between the fact that *p* and the belief that *p* if that belief is to constitute knowledge (a connection that she endorses and for which she argues). This connection is lacking if one's belief that one's ticket will lose is based on merely probabilistic grounds—it is missing in the lottery case, and that is why one does not know that one's ticket will lose. Secondly, she postulates that justification requires an internalized version of the causal or explanatory connection that knowledge requires (which is not to say that that is *all* that justification requires). To be justified in believing that *p*, the believer must be able to "suppose that there is . . . a causal or . . . explanatory connection between one's belief and its object" (Nelkin, 2000, p. 397), and Jim cannot see that there is any such connection for his belief that his ticket will lose—indeed, he can see that there is no such connection. Nelkin notes that the believer's supposition that there is a causal or explanatory connection between his belief that *p* and



the fact that  $p$  must itself be justified if that supposition is to render the belief that  $p$  justified.<sup>33</sup>

I do not see that Nelkin's introduction of causal and explanatory relations is of any utility except insofar as it explains why Jim does not know that his ticket will lose (a utility that Nelkin explicitly takes it to have). If we take it for granted that Jim does not know that for *whatever* reason, it seems that talk of causal and explanatory relations adds nothing to Nelkin's account of why Jim's belief would not be justified. The relevance that Jim's inability to see such connections has to his lack of justification is just that it renders Jim unable to postulate justifiably that he *knows* that his ticket will lose, the connections in question being precisely those that are necessary for a belief to constitute knowledge. So, what Nelkin's account of why (1\*) is false boils down to is that a belief that  $p$  is not justified unless the believer justifiably believes (or *could* justifiably believe or would justifiably believe upon reflection) that he knows that  $p$ . Now, Nelkin does not explain *why* justified belief should require justifiably believing (at least on reflection) that one knows that  $p$ . Indeed, since she stresses that Jim can "see" (her term) that his belief does *not* have the appropriate causal or explanatory connections to its object, all she is really doing is pointing out that Jim is in a position to know that he does not know that his ticket will lose. That is, she is not explaining why premise (1\*) is false, but simply pointing out that lottery beliefs belong to that category of allegedly justified beliefs we have been calling the known unknown beliefs. Nelkin is doing little beyond baldly asserting that such beliefs are not justified.<sup>34</sup> Her "explanation" is the start of a genuine explanation of the falsity of (1\*), however. The reason *knowledge* is at the heart of her characterization of justification is that justification just *is* knowledge, and the lottery beliefs are hence unjustified since they do not constitute knowledge.

(What might a proponent of justification as warrant say in response to the lottery paradox? Since what is characteristic about warrant is that a warranted true belief is supposed necessarily to be knowledge, the only tenable strategy is to endorse Nelkin's claim and my own that (1\*) is false. If one's belief that one's ticket will lose is warranted, then, provided one's ticket actually does lose, one knows that one's ticket will lose; this is not so, hence we must conclude that one's belief that one's ticket will lose is not warranted. The same reasoning renders the rest of the class of known unknown allegedly justified beliefs unwarranted, unless one wants to claim that only the false examples are warranted! (That having been said, I am not at all sure how many popular accounts of warrant avoid a commitment to our knowing that our tickets will lose. A general strategy of forming such probabilistically based beliefs certainly reliably yields truths; it is not clear why properly functioning faculties designed to yield true beliefs would not produce such beliefs, and so on.))

### 4.3 The Modesty Argument

This argument concludes that there are no known unknown justified beliefs. It is the only argument whose conclusion concerns only *some* of the allegedly justified beliefs. Although my other arguments make use of a proper subset of the allegedly justified beliefs (e. g., the lottery argument), their conclusions are general: the only justified beliefs are those that constitute knowledge.

If the belief that *p* is one of the known unknown beliefs, then, by definition, one is generally in a position to know that one does not know that *p*. I am not, of course, in any way claiming that one is infallible in determining that one does not know that *p*. No doubt thinkers fairly often take themselves to know theories that are false or, at best, lacking sufficient support to constitute knowledge. No doubt thinkers often take themselves to know propositions that at best they know to be probable. Most such thinkers (at least) have some unjustified beliefs, however—their estimations of the extent of their knowledge. I am arguing that those thinkers and others who do not explicitly overestimate what they know—who do not believe that they know propositions that they do not know—but still have beliefs in the known unknown class are unjustified in *those* beliefs, rather than their beliefs about the extent of their knowledge.<sup>35</sup>

The skeleton of the argument is suggested by Nelkin in her attempt to explain away the appeal of the claim that Jim is justified in thinking that his lottery ticket will lose ('the lottery claim'):

Strongly associated with the appeal of [the lottery claim] is the conviction that Jim should not make plans contingent on his winning the lottery, and, in fact, that he should make plans based on the assumption that he is not going to win. Similarly, we think that he should not "get his hopes up," that he should not devote excessive energy to creating mental images of his winning, and that [perhaps] he should not buy a lottery ticket . . . In fact, I believe that the appeal of [the lottery claim] derives largely from the fact that Jim's belief that he will lose rationalizes all of these behaviors that we find rational and normative for Jim. But . . . compatible with the falsity of [the lottery claim] is the truth of

(1\*\*) Jim is [justified] in believing that he will very probably lose.

All of the behaviors just mentioned (making plans on the assumption that he will lose, not getting his hopes up, and so on) are sufficiently rationalized by Jim's belief that he will very probably lose. (Nelkin, 2000, p. 400)

Explicitly probabilistic beliefs such as that described in (1\*\*) are justified on anyone's view, including my own since they constitute knowledge in the relevant cases. Nelkin suggests that these beliefs are at least as good a guide to behavior as their bolder, non-probabilistic counterparts from a third-person perspective *and*, more importantly, from a first-person perspective. One gains nothing—no motive for rational action, and no benefits thereof—

from believing immodestly that one will lose the lottery that one does not gain from believing modestly that one will very likely lose the lottery.

This is not to say that believing that one is very likely to lose the lottery and believing that one will lose the lottery do not differ at all in their (potential) effects on a believer's behavior. Believing categorically that one will lose the lottery will lead to the kind of behavior that is associated with greater confidence than its probabilistic counterpart. But such behavior will be the behavior of the irrationally overconfident. Indeed, if one believes categorically that one's ticket will lose the lottery, one invites the question that is often posed rhetorically to illustrate that one does not *know* that one's ticket will lose. If one believes categorically that one's ticket will lose, why did one buy the ticket in the first place? (That this does seem like an appropriate question is nicely explained by my view that belief *aims* at knowledge, that a belief is only properly held when it constitutes knowledge.)<sup>36</sup>

So much for the probabilistic subdivision of the known unknown beliefs—what about what we might term the “theoretical” subdivision, beliefs in theories that are best supported by the evidence, but not so well supported that one knows them to be true? What is the modest thing to believe which will gain one all the advantages as a guide to rational action that a brasher belief would? Simply that the evidence supports one theory over the others (and to such-and-such a degree, insofar as one can make such a claim). Believing in the theory outright will, once again, be an instance of irrational overconfidence, and invites a question similar to one that arose in the probabilistic case, at least for “active” proponents of a theory rather than those who believe the theory because of, say, what they read in the newspaper. If one categorically believes the theory to be true, why is one still investigating its truth—gathering further evidence (if one is a scientist concerned with an empirical theory), looking for new arguments (if one is a philosopher concerned with a philosophical theory), and so on?<sup>37</sup>

Why not be maximally modest, one might ask? Do not *ever* believe that a theory is true, simply believe that the evidence supports (or *appears* to support?) it more or less well even if one does know it to be true—or would if one so believed. Never believe that one has lost the lottery, simply that it is very likely that one did, even when one watched the draw take place and knows that one lost (or would if one so believed). Indeed, never believe that one has hands, simply that it seems to be so. Does not our previous discussion lead to an endorsement of this absurd (and, no doubt, psychologically impossible) cognitive behavior?

It does not. The absence or presence of knowledge makes the difference between rational and irrational belief formation. The vast majority of discussions of irrational belief and irrational believers focus on beliefs that believers form that they should not. Some irrational believers qualify as such because of the beliefs that they *fail* to form, however. Some

non-philosophical sceptics (whose scepticism is restricted to specific areas of enquiry) are too sceptical. They should form beliefs that they do not form. Often (at least), they should form those beliefs because the non-sceptics who believe the contested propositions *know* them to be true, and on the same grounds that the sceptic himself possesses.

Alternatively, consider someone suffering from an obsessive-compulsive disorder who just will not stop washing his hands. He does not think that they are clean, but he has such strong grounds for thinking that his hands are clean that if he were only to believe, he would know. He *might* further believe that his hands are not clean, an irrational belief of an orthodox kind. But even if he were *agnostic* about the state of his hands, his handwashing would be irrational because motivated by an irrational *lack* of belief. (From the point of view of desire or goal satisfaction, his behavior is apparently unobjectionable—he does not believe that his hands are clean, and he wants them to be clean, so he is washing them.) The maximally modest believer would suffer from the same kind of irrationality—he refuses to believe what he is in a position to know. Moreover, restricting oneself *per impossibile* to believing merely that one *seems* to have hands would, I suppose, result in behavior characteristic of cognitive underconfidence. Would one attempt to block the rock hurtling toward one's head rather than throw oneself on one's back if one merely believed that one seemed to have hands?

An optimal believer, whose beliefs are precisely as modest as is necessary to produce optimal behavior, neither overconfident nor underconfident in belief or subsequent action, is one who restricts his beliefs to what he knows. Such belief accords with primary epistemic goals, among which is, arguably, the optimal guidance of action—it is, then, justified in an evaluative sense. Being an epistemic ideal, such modesty accords well with a Kornblithian understanding of deontological justification; one ought to be ideally modest, which requires believing only what one knows. (And avoiding the knowledge-shy failure to form beliefs noted above. There is at least one positive epistemic obligation: if one would know that *p* were one to believe that *p* simply in virtue of so believing, one ought to believe that *p*.)

#### 4.4 The Posterior Evaluation Argument

If a belief that *p* is one that would be justified were one to form it, and it is in one's interest to have a belief in whether or not *p*, and one is capable of forming such a belief, then, in some intuitive sense, one should believe that *p*, although that 'should' is not purely a 'should' of epistemic obligation. It is rather a 'should' generated by prudential considerations of self-interest interacting with the epistemic goals that determine an evaluative notion of justification.<sup>38</sup> This claim is clearly true if we read 'would be justified' as 'would constitute knowledge'. I will argue that it is not true on any more expansive conception of evaluative justification that incorporates the unknown unknown class of beliefs, casting doubt on such conceptions.

Whatever justification is, it is surely *valuable*, even if it is just one of several epistemic values. If there can be justification without knowledge, when knowledge is absent, but justification is present, something of value remains, at least if no factor *independent* of knowledge is present to remove the value of justification—a factor, that is, that might or might not be present when knowledge is absent. The situation described below is explicitly constructed to contain no such factor. The mere absence of knowledge cannot remove the value that justification has—unless, of course, justification is knowledge. The opening claim that one should in *a* sense believe that *p* in the specified conditions is compatible with there being other senses in which one should *not* believe that *p*, or even other senses in which one should believe that not-*p*. There are beliefs that certain individuals should not form in an important sense because the beliefs would be dangerous, upsetting, or even lethal—even if they would constitute *knowledge* if the individuals were to form them. I stipulate that the situation described below is not disposed to produce such beliefs. (As noted in section 3.1, no members of the class of unknown unknown beliefs are warranted in Plantinga's sense, not even the false ones, and hence the argument of this section is not an argument against any notion of warrant.)

Suppose that one has never driven a Ford, and one wants to. We might even suppose that there is some financial (or romantic, or whatever) advantage to one's doing so. And let us suppose that one has plenty of evidence that one has a (particular) colleague who has a Ford, and one knows that all of one's colleagues would let one drive their cars. However, that evidence is misleading. As in the famous Gettier case, one does have a colleague who owns a Ford, but it is not the colleague that one's misleading evidence indicates owns a Ford. Were one to form a belief on the basis of one's evidence that one has a colleague who owns a Ford, that belief would be true. Let us suppose further that it would be justified despite not constituting knowledge.

However, one fails to notice the apparent implications of one's evidence. One fails to form any belief on the basis of that evidence, and so does not come to the justified conclusion that one has a colleague who owns a Ford. Some time later, one realizes that one previously failed to realize the apparent implications of one's evidence, and one is aware that had one done so, one would have formed the belief that one had a colleague who owned a Ford. Further information has since come to light concerning the misleading nature of one's former evidence, however, and one is now aware that the belief that one would have formed would have failed to constitute knowledge despite being true. At that later time, then, one does know that one had a colleague who owned a Ford, although, let us suppose, it is now too late to garner the benefits that would have accrued to one had one believed earlier that one had such a colleague. Is there an intuitive sense in which one will now judge that one should then have formed the belief that

one had a colleague who owned a Ford—an allegedly justified, true belief? Clearly, there is. That belief would have been true.<sup>39</sup> So one should have formed it. However, we can say exactly the same about a similar situation in which one has no evidence whatsoever that one has a colleague who owns a Ford, although in fact one has such a colleague. There is an intuitive sense in which one should have formed the belief that one had such a colleague despite the fact that such a belief would have been unjustified by anyone's lights. Such a belief would have been true.

Consider also a situation that is similar except insofar as one's very good evidence that one has a colleague who owns a Ford is *not* misleading; it is quite sufficient for its possessor to gain knowledge that he has a colleague who drives a Ford (the very colleague to whom the evidence points), provided he forms that belief. Which, once again, one fails to do. Evaluating one's former cognitive situation after the fact, is there an intuitive sense in which one should have formed the belief *apart* from the consideration that it would have been true? I contend that there is: one should have *known* that one had a colleague who owned a Ford, and this is to say, in the case at hand, that one has let oneself down doxastically speaking in a way that goes beyond having merely failed to form a true belief (for which one could not coherently blame oneself, the beneficial consequences notwithstanding).<sup>40</sup>

So, in one case there is an intuitive (but non-epistemic) sense in which one should have formed the belief that one had a colleague who owned a Ford because it would have been true, and in a different case a different intuitive sense (one partially epistemic) in which one should have formed that belief because it would have constituted knowledge. What we lack, I claim, is the intermediate case. When one could have formed an allegedly justified, true belief that one had a colleague who owned a Ford which fell short of knowledge but one failed to do so, there is no intuitive sense in which one should have done so beyond the non-epistemic sense in which one should have done so because it would have been true (and beneficial). We should expect there to be such a sense, a sense in which one should have formed a justified belief on the question of whether or not one had a colleague who owned a Ford given that it is in one's interest to do so and that one was capable of so doing. The absence of the intermediate case is explicable if the sense in which one should have formed a justified belief on the matter in question just *is* the sense in which one should have known that one had a colleague who owned a Ford since justification is knowledge. A similar case can be made for any belief in the unknown unknown class of allegedly justified beliefs; none of them is genuinely justified.

## 5. The Main Objection

How one responds to objections to one's view and to one's arguments for one's view clarifies the view and the arguments for the view themselves, particularly when one's view is as unorthodox as my own. Consequently, I have explained many objections and my responses to them in the course of laying out my view and my arguments. One objection is so potentially devastating that it deserves separate treatment. If successful, the objection demolishes three of my four arguments; all but the posterior evaluation argument, which is arguably the least intuitively compelling of my arguments in any case. A view as unorthodox as mine demands more than a single argument. In this section I will explain the main objection to my arguments, and how it threatens to defeat three of my arguments. I will then argue that the main objection is unsuccessful.

The main objection is very similar to an objection that Williamson discusses to his knowledge rule for assertion; it is, in a sense, an upgraded version of that objection which aims to avoid one of his counterobjections. I will argue that his counterobjection can also be upgraded to render the objection powerless once more. The objection is that a weaker norm governing assertion will account for the phenomena that Williamson cites in support of his knowledge rule. That rule is:

(The RBK rule) One must: assert  $p$  only if one rationally believes that one knows  $p$ .  
(Williamson, 2000, p. 261)

The RBK rule accounts for the quasi-Moorean phenomena. One cannot assert a sentence of the form ' $p$  and I do not know  $p$ ' since if such an assertion were correct, one would rationally believe both that one knows  $p$  and that one knows that one does not know  $p$ . If one rationally believes that one knows that one does not know  $p$ , then one rationally believes that one does not know  $p$  since knowledge entails truth and one knows that. So one rationally believes that one knows  $p$  and that one does not know  $p$ . That is impossible; the quasi-Moorean assertion is unwarranted on the RBK rule. The RBK rule also accounts for the fact that lottery assertions are unwarranted. One cannot assert that one has not won the lottery since one cannot rationally believe that one knows that one has not won the lottery. The rule also explains why 'How do you know?' and related statements are appropriate challenges to many assertions; a challenged asserter should be able to explain how he rationally believes that he knows what he asserted, which is to explain how he knows what he asserted if the asserter actually does know what he asserted.

The first of Williamson's objections to the RBK rule is, fundamentally, that it makes assertions that express false unknown unknown beliefs warranted:

Suppose that I rationally believe myself to know that there is snow outside; in fact, there is no snow outside . . . yet something is wrong with my assertion;



[the RBK account does not imply that there is]. [It] can allow that there is something wrong with my belief that I know that there is snow outside . . . The [RBK account] lacks the resources to explain why we regard the false assertion itself, not just the asserter, as faulty.

(Williamson, 2000, p. 262)

The defender of the RBK account (we will stick with the name only temporarily) should say, I think, that this is not a *consequence* of his view—it *is* his view. The view is that it is not just assertions that express knowledge that are warranted, but assertions that express any unknown unknown belief, whether true or false. Indeed, one does not know that one does not know that *p* (even upon minimal reflection) if one has an unknown unknown justified belief that *p* by definition; perhaps if one believed that one did *not* know that *p*, then one would not be justified in believing *p*.<sup>41</sup> That does not require that one believe that one does know that *p*. One might not have considered the matter, and one's belief and any assertion that expresses it are not *ipso facto* unwarranted. It is time to upgrade the RBK rule:

(The J rule) One must: assert *p* only if one has a justified belief that *p*.

Justified beliefs are those that constitute knowledge, together with the unknown unknown beliefs. The known unknown beliefs are simply not justified; that is the conclusion that we should draw from the assertion phenomena that Williamson uses to argue for the knowledge rule, which the upgraded J rule explains as well as the RBK rule, and in the same way. If one has a justified belief that *p*, then one cannot also have a justified belief that one does not know that *p*. The unknown unknown beliefs are would-be knowledge; in unexceptional circumstances, where one does not suffer from bad luck in one's external environment, one would know what one justifiably believes. Clearly one cannot both know that *p* and know that one does not know that *p*, and neither could one have done so but for exceptional bad luck. The quasi-Moorean phenomena are explained by the J rule. Lottery beliefs are not justified, so neither are assertions that express such beliefs. If one has an unknown unknown justified belief that *p*, since one would not know that one did not know that *p*, one will, upon reflection at least, believe that one does know that *p*; if one does not, one would either not have been justified in one's belief in the first place, or would cease to be so upon reflection. Otherwise, it is hard to see how one would have known that *p* but for exceptional bad luck in one's external environment. Consequently, the 'How do you know?' challenge to a warranted assertion is explicable even for those warranted assertions that express unknown unknown beliefs rather than knowledge.<sup>42</sup>

The J rule does indeed imply that certain false assertions are warranted, and Williamson claims that falsity is a defect in an assertion over and above

the defect that it is in any belief that an assertion expresses. I find that claim considerably less intuitive than his related claim that the flagrant irrationality of an assertion is a defect over and above the defect that it is in a flagrantly irrational belief so expressed—a claim that he makes to dispatch the weaker BK rule that one must assert only what one believes that one knows (rationally or irrationally). Besides, falsity is not a defect in a belief in the sense that it renders the belief unjustified on the *standard* view of justification; it would be puzzling if falsity rendered an assertive expression of a justified belief unjustified. (Of course, it does no such thing on my view: a false belief is unjustified *per se* since it does not constitute knowledge.)

A second objection to the RBK rule that Williamson makes is that the knowledge rule provides a *simpler* account of warranted assertion than the RBK rule, and so the burden of proof is on the proponent of the RBK rule to refute the account based on the knowledge rule. It is here that upgrading the RBK rule to the J rule pays off. The RBK rule is indeed more complex than the knowledge rule as its three-letter acronym clearly reveals. The J rule, on the other hand, dispenses with the B and the K components of the RBK rule.<sup>43</sup> It is not obviously any more complex than the K rule, and so no account of warranted assertion based upon it clearly has the burden of proof.

I will argue that Williamson's complexity argument against an RBK account has a counterpart that will defeat an account based on the J rule. Before giving the counterpart argument, let us see how a proponent of the J rule can defeat three of my arguments. It is obvious how he can defeat the assertion argument. That Williamson's knowledge rule governs assertion was a premise of that argument; if it is rather the J rule that governs assertion, then that premise is false.

It is almost truistic to say that the J rule (or a close counterpart) is as much a rule that governs warranted belief as warranted assertion. One must not believe that *p* unless one is justified in so doing. The rule has some nontruistic bite when its proponent adds that it is only the unknown unknown beliefs along with those beliefs that constitute knowledge that are justified; the known unknown are unjustified. Let us call the relevant notion of justification 'U-justification'. The modesty argument concludes that the known unknown beliefs are not justified. It is hence not a problem for the view that a belief is justified iff it is U-justified. As Nelkin herself in effect argues, we can give uniform solutions to the knowledge and justification versions of the lottery paradox by denying that known unknown beliefs, such as beliefs that one will not win the lottery, are justified; *why* they are unjustified is a further question to which the proponent of U-justification need not supply an answer. The tenability of his view does not obviously hang upon the provision of an answer.

In any case, perhaps pointing out that one can know that one does not know that one will not win the lottery is enough of an answer. I *do* have an

explanation of why the lottery belief is unjustified: justification is knowledge. Is that a better explanation than: justification is U-justification? I suggest that it is, *especially* in light of the resurrection of Williamson's complexity argument below. Without that resurrection, perhaps the proponent of U-justification could *half*-defeat the lottery argument. A full defeat—the provision of an equally compelling response to both versions of the lottery paradox—would require an explanation of why justification is U-justification.

Now for the resurrection of Williamson's complexity argument. The J rule is, in fact, more complex than the knowledge rule because of the considerations of section 2.1. U-justification is more complex than knowledge since the concept of U-justification is parasitic on that of knowledge. A U-justified belief is a belief that constitutes knowledge or a belief that would have done so but for bad luck in the believer's external environment. At the very least, the burden of proof is shifted back to the proponent of the J rule. He must show that U-justification is in fact a more primitive concept than that of knowledge rather than one that we understand as a disjunction of knowledge and would-be knowledge.<sup>44</sup>

## 6. Conclusion or The Bit Where You Take It Back, Part II

On anyone's view, people form a lot of unjustified beliefs. On my view, they do so even more commonly than on perhaps any view other than a philosophical sceptic's. In this section, I want to suggest that unjustified beliefs, although common on my view, are not quite as common as one might suppose since belief—or at least belief within the known unknown class of allegedly justified beliefs—is less common than one might suppose.

Utterances of the form 'I believe that *p*' and similar forms ('I think that *p*', '*p*, I believe', 'I think so', etc.) often, I suggest, do not express belief in the proposition that *p*. They express, rather, a belief that *p* is probable (more likely than not, perhaps). Consider a stranger who asks one where the post office is. One does not know, but has a vague idea that it is a mile to the right. Consequently, one says 'I believe that it is a mile down that way'. This is a perfectly proper utterance; I suggest that it is also not literally true. One does not believe categorically that the post office is a mile to the right; one believes that it is more likely there than not. Ironically, a sentence of the form 'I believe that *p*' is sometimes used to convey precisely that one does *not* believe that *p*, strictly speaking.<sup>45</sup> Even if I make an explicitly probabilistic self-ascription of belief, such as 'I believe that it is more likely than not that there are mice in the basement', it is entirely proper to report my belief without a probabilistic modifier—'He thinks that there are mice in the basement'.<sup>46</sup> Third-person belief ascriptions need not attribute belief in the

proposition that an overly literal interpretation would suggest any more than their first-person counterparts.

Scientific theories provide another illustration. I take it that 'It is believed that birds are descended from dinosaurs' is a perfectly proper statement (or, if it is not, it is down to my ignorance of the debate over the relation between birds and dinosaurs). This is so even if *no one* literally believes that birds are descended from dinosaurs, but the strongest view that anyone holds is that they probably did, or that the available evidence supports that theory better than any rival theory, or some such. (Perhaps phenomena like these prompt talk among some philosophers of theory *acceptance* in science rather than belief therein, although some such philosophers seem to take acceptance to be less easily reducible to belief than a notion of acceptance based on the considerations above.)

Since the qualified beliefs that I suggest apparently bolder ascriptions of belief in fact ascribe do, in general, constitute knowledge for their holders, unjustified belief is a lot less common than an overly literal understanding of statements of belief would indicate.

Alvin Goldman (Goldman, 1999) associates a deontological conception of justification with the notion of *guidance*; epistemic duties, if there are any, are supposed to play a role in how one forms one's beliefs. The title of this paper indicates how an identification of justification with knowledge enables the believer to be so guided. Although we perhaps do not always know that we know what we know, and perhaps sometimes *cannot* know that we know, we nevertheless are *often* able to know that we know—or that we do not know. If we find ourselves believing theories that we know we do not know to be true, although we take ourselves to know them to be best supported by the available evidence, we can retreat to the more modest belief. If we find ourselves believing that our house has not burned down while we were on vacation, we can correct ourselves and believe that it is merely very *probably* intact. If we find ourselves believing everything we read in the newspaper, although we know that identifiable portions of it are unlikely to express knowledge on the part of the reporter (and so beliefs thereby produced will fail to constitute knowledge), we can retreat to what we *do* know, which might be as much as that there is a good chance that what was reported is the case, or as little as that it was reported to be.

Readers might wonder if I really believe the view defended here. I assure them that I do. That, it should be clear by now, is a very strong commitment indeed.<sup>47</sup>

## Notes

<sup>1</sup> Linda Zagzebski is one philosopher who does stress the bad luck involved in the Gettier cases (Zagzebski, 1996, p. 284).

<sup>2</sup> See Heller (1999) for one instance of this quite common claim.

<sup>3</sup> Crispin Wright makes just such a suggestion (Wright, 1991); the oral expression of the sentiment is common.

One might think justification more fundamental than knowledge simply because one wants a notion of justification in terms of which one can *define* knowledge. For more on this matter, see section 3.1 below.

<sup>4</sup> We cannot characterize the unknown unknown beliefs simply as those that, although they do not constitute knowledge, are based on good reasons. This is an insufficiently discriminating characterization. Many philosophers will take a belief that one's ticket will not win a fair lottery to be based on good reasons, but such lottery beliefs are not unknown unknown beliefs. They would not amount to knowledge that one's ticket will not win even in normal circumstances, and so they cannot easily be used to construct Gettier cases. Lottery beliefs are members of the second category of allegedly justified but unknown beliefs discussed immediately below—the known unknown.

If by 'based on good reasons' one means 'based on reasons sufficient to render a belief knowledge in normal circumstances', or if one denies that a belief that one's ticket will not win the lottery is based on good reasons since it would not constitute knowledge even in normal circumstances, then I do not have any quibble with the characterization of the unknown unknown beliefs as beliefs based on good reasons that do not constitute knowledge—because it does not clearly differ from my own more than terminologically. The claim that labelling the unknown unknown beliefs 'justified' employs a concept of justification that is parasitic on knowledge stands.

<sup>5</sup> A referee suggests that if the believer realizes that he does not know that *p* upon reflection, this will be a defeater for his allegedly justified belief that *p*. The view that he *would* be justified if and only if he did not come to realize that he does not know that *p* is only appealing, I suggest, if one ties justified belief very closely to blameless belief, a position that I argue should be rejected in section 3.4. The believer might well be blameless in failing to realize that he does not know that *p* and for that very reason in his belief that *p* itself.

The view that the believer will not have a justified belief that *p*, whether or not he realizes that he does not know that *p* in large part because he could come to realize that he does not know that *p* upon minimal reflection is in accord with my own view that the known unknown beliefs are not justified, and so I do not take exception to it.

<sup>6</sup> Zagzebski and Merricks argue for this claim on similar grounds: any truth-independent notion of warrant used to define knowledge will render the definition open to Gettier-like counterexamples.

<sup>7</sup> I agree with Plantinga and others that the intuitiveness of more orthodox notions of justification arises in large part from a conflation of blamelessness and justification (about which I shall talk more in section 3.4).

<sup>8</sup> I use the term 'obligation' simply to denote deontic facts expressed by such statements. This is somewhat at odds with ordinary usage. Perhaps I should not have believed that Santa Claus would visit on Christmas Eve when I was a child, but it is exceptionally odd to say that I had an obligation not to believe that. In the ordinary sense of 'obligation', its connection to notions such as blameworthiness might be unbreakable. That connection is breakable if we use the term for the deontic facts alone, as I argue in section 3.4 below.

<sup>9</sup> Feldman no longer endorses the stronger claim (Feldman, 2000), instead suggesting that, if one adopts an attitude towards a proposition at all, it should be the attitude (and, he says, there is one and only one, understanding withholding belief and disbelief as attitudes) that one's evidence supports. One is not, on Feldman's later view, generally obligated on purely epistemic grounds to adopt an attitude towards a proposition.

<sup>10</sup> At least if the claim is that belief aims at truth *simpliciter*, rather than something that at least *entails* truth.

<sup>11</sup> Williamson (2000, ch. 1) makes just such a claim.

<sup>12</sup> I expect some internalist dissent here, but here is not the place to debate the merits of internalism about justification.

<sup>13</sup> If there are several primary epistemic goals, then they might conflict in an individual case—one fulfils one goal by forming a given belief, but *not* forming the belief fulfils another. In such a case, it is conceivable that one ought not form the belief. What I am arguing is that if forming a belief satisfies one or many primary epistemic goals, and not forming it does not satisfy any primary epistemic goal, then it is inconceivable that one nevertheless ought not form the belief.

<sup>14</sup> I trust I do not need to define precisely what I mean by ‘relevant mental state’ here; what I am saying should be clear enough, and fairly uncontroversial. By ‘relevant’, I do not mean to imply ‘all’, in part because Williamson argues that knowledge itself is a mental state (Williamson, 2000).

<sup>15</sup> His mental states differing only insofar as the change in external circumstances entails such a difference—as it would on Williamson’s view mentioned in the previous footnote.

<sup>16</sup> One is tempted to say that no matter how well one knows one’s own mind, there is almost no proposition about the external world that one is thereby guaranteed to know. But again, that characterization threatens to conflict with Williamson’s views.

<sup>17</sup> Kornblith makes these observations primarily to argue that voluntarism about belief is not entailed by the existence of epistemic obligations.

<sup>18</sup> Although many properties that some beliefs have are objectively bad from the perspective of truth-conduciveness, making his non-factualism only partial.

<sup>19</sup> However, an internalist has no obvious response to my arguments below that justification is knowledge.

<sup>20</sup> Plantinga and other proponents of warrant make similar claims.

<sup>21</sup> “Warrant” is Williamson’s term; Unger, and DeRose following him, say that one “represents oneself” as knowing that *p* in so asserting. Williamson’s use of the term is not to be confused with Plantinga’s technical notion, of course.

<sup>22</sup> I am, of course, supposing that it is possible to gain justified beliefs and knowledge through testimony, an assumption that should need no defense nowadays.

<sup>23</sup> Perhaps any theory of justification that allowed the recipient of testimony to have a more highly justified belief acquired solely on the basis of testimony than the testifier himself would be *ipso facto* objectionable. Perhaps not, however—the testifier is acting as a filter to make the trustworthiness of his assertions exceed that of his raw beliefs, and that is not obviously an absurd notion.

<sup>24</sup> It is even possible that Bob knows that Andy does *not* know (Andy might tell him that he does not know, for example), but is merely justified (assuming such a thing is possible for the sake of argument, of course), if Andy’s justified belief is a member of the known unknown class. (If it is a member of the unknown unknown class, just as Andy could not come to know that he did not know without also losing justification for his belief, Bob could not come to know that Andy was merely justified in believing that *p* and acquire justification for believing that *p* through knowing that about Andy.) Bob still ought to be able to acquire a justified, true belief on the basis of Andy’s testimony.

<sup>25</sup> I say that the goal is to transmit beliefs rather than justified beliefs. This does not entail that it is also a goal of assertion to transmit *false* or *unjustified* beliefs. I might have the goal of learning to play the piano without having the goal of learning to play the piano well. This is not to say that I also have as my goal to play the piano badly.

<sup>26</sup> We suppose here that the notion of warrant can be applied univocally to both assertion and belief. If ‘warrant’ is *just* employed as a term for a component in the definition of knowledge which is not alleged to have any independent status (that is, it is not also a deontological or evaluative notion of justification) then my arguments do not touch it, nor are they intended to. I doubt knowledge can be defined in terms of such a notion, but arguing the case is not my concern here.

<sup>27</sup> If definite descriptions and possessive pronouns are also quantifiers, the actual truth communicated might be considerably more verbose, but I trust that the idea is clear.

<sup>28</sup> I am not the first to propose that Williamson's knowledge rule governing assertion entails a corresponding knowledge rule governing belief. Crispin Wright (Wright, 1996, p. 935) takes this line on the grounds that "the . . . notion of warranted assertion . . . is simply the exterior counterpart of warranted belief." (This sentence is taken from a single paragraph concerning the knowledge rule—he does not give any further argument, nor can he be expected to given the context in which his discussion occurs.) Wright, however, takes the entailment to be grounds for rejecting Williamson's knowledge rule governing assertion since "[the] long tradition in epistemology of distinguishing knowledge and reasonable belief may indeed be misguided; but, if it is, it will demand a very substantial argument to show it." I hope this paper will provide such an argument. (Wright also notes that assertions made in many contexts seem to fall short of expressing knowledge, contexts such as "medical diagnosis, weather forecasting, ordinary psychology, history, economics [and] plant ecology," and he seems to think that a presumption of warranted assertion in those contexts outweighs the intuition that Williamson's lottery assertions are unwarranted. I suggest that ordinary politics and ordinary journalism should be added to Wright's list.)

<sup>29</sup> Actually, she presents the paradoxes as concerning knowledge on the one hand and *rationality* on the other. However, it is clear that the rationality paradox can be restated without loss in terms of justification, and footnote 13 of Nelkin's paper indicates that she is happy with such a restatement provided that justification is not understood as warrant in Plantinga's sense. I will briefly mention how a proponent of justification as warrant might approach the paradoxes below.

<sup>30</sup> Nelkin's locution "It is rational for Jim to believe" does not entail that Jim actually believes, and I am keeping this feature of the argument intact in my representation of it in terms of justification. I intend to be expressing the thought that if Jim does not believe that his ticket will lose, then if he were to do so, his belief would be justified, and if he does so believe, then his belief is justified.

<sup>31</sup> If we employ a deontological phrase, to say that it is permissible for Jim to believe that his ticket will lose is to say that if he were to so believe, then he would know. His belief would be in accord with the epistemic duty to believe only if one knows.

<sup>32</sup> She reviews and finds wanting a number of alternative actual and potential joint diagnoses of the two versions of the paradox (although not my own).

<sup>33</sup> 'Rational' being the term that she employs, as previously noted. Obviously, this condition on justified belief is not supposed to form part of a definition of justification because of any such definition's obvious circularity.

<sup>34</sup> Although perhaps she means to assert this of the probabilistic examples alone, theoretical beliefs arising from inferences to the best explanation that fail to constitute knowledge having sufficient explanatory connection to their objects to be justified. Her text is unclear on the point since her example of a general belief justified by (a disposition to) justified belief in explanatory connections that are not causal is the belief that all objects on the Earth are affected by its gravitational pull. This, of course, is something that we know and so is not a non-probabilistic member of the known unknown class of allegedly justified beliefs. It remains the case, then, that the examples that she provides of causal and explanatory connections (potential) belief in which is supposed to be essential to justification are precisely those connections that support knowledge. She gives no indication that she would accept weaker connections than that; if she did, it would not *ipso facto* supply her rejection of (1\*) with a genuine explanation of its falsity in any case.

<sup>35</sup> In a sense, of course, *any* belief that falls short of knowledge implicitly overestimates the extent of one's knowledge on my view.

<sup>36</sup> One might object that categorical belief that one will lose is entirely appropriate, but the *degree* of that belief should be less than maximal. Talk about degrees of belief can be



interpreted in terms of simple belief in explicitly (which is not to say *precise*) probabilistic propositions, on which interpretation, the “objection” to my view is no objection at all—it is my view. Although there is a tradition of talking of degrees of belief that does not seem consistent with such an interpretation (and does not amount to a measure of how willingly one would give up one’s belief, which is irrelevant here), I will simply register my suspicion that it is far from clear that such talk is comprehensible. The topic is too large to deal with at length.

<sup>37</sup> To say that we recognize our own fallibility is not enough to answer the challenge. We also recognize that we must terminate enquiry at some principled point despite our fallibility. Categorical belief, governed by the goal of knowledge, provides such a point.

<sup>38</sup> If the principle is restricted to beliefs that would be true if formed, the argument to follow will still proceed.

<sup>39</sup> And, *ex hypothesi*, a true belief on the matter would have had beneficial consequences. I stipulate this since I do not wish to assume that a true belief itself is a good even if it has neutral or even harmful consequences. If a true belief *does* have beneficial consequences, one should, in some sense, form it *in addition* to the fact that one should experience those consequences (quite apart from the fact that they are consequences of a true belief).

<sup>40</sup> I emphasize once again that one’s obligations and subsequent blameworthiness are generated by a combination of epistemic considerations and prudential and other non-epistemic considerations particular to one’s situation. One is not in general obligated to know what one *could* know, and neither, as we stressed in section 3.4, does blameworthiness follow from unmet obligation in all or most cases. In some cases these phenomena go together, and we can suppose that the proposed situation is one of these cases.

<sup>41</sup> The optimal defense of the RBK account endorses just such a possibility as I explain immediately below.

<sup>42</sup> If one interprets the ‘R’ of the RBK rule along the lines of the ‘J’ in the J rule, the RBK rule is arguably *stronger* than the knowledge rule itself. A justified belief that one knows that *p* will, in unexceptional circumstances where one is not subject to bad luck in one’s external environment, constitute knowledge that one knows that *p*. Such an RBK rule is not as strong as a KK rule that *requires* one to know that one knows what one asserts, but, in unexceptional circumstances, asserters who follow an RBK rule thus interpreted will know that they know what they assert.

<sup>43</sup> Of course, *belief* is still a component of the J rule, just not belief that one knows; it is no less a component of the knowledge rule in any important sense, since knowledge entails belief.

<sup>44</sup> An account of *warrant* (section 3.1) will not meet the burden of proof since such an account, if successful, *avoids* the Gettier problem by classifying U-justified true beliefs as nevertheless unwarranted. There is no reason to think a successful definition of warrant will provide a basis for a successful non-disjunctive definition of U-justification.

<sup>45</sup> Another view is that the belief sentence *is* true, rather than being a falsehood used to communicate a truth. It is just that sentences apparently concerning belief in one proposition are sometimes made true by belief in a different but related proposition. It is not necessary to decide between these (and other) options in the philosophy of language for our purposes.

<sup>46</sup> Contrast with the unacceptability of reporting my assertion ‘There are mice in the basement, more likely than not’ with ‘He says that there are mice in the basement’.

<sup>47</sup> I believe that the *details* merely have an even chance of being correct.

## References

- Alston, W. P. (1985), “Concepts of Epistemic Justification,” *The Monist*, 68:57–89  
 Alston, W. P. (1993), “Epistemic Desiderata,” *Philosophy and Phenomenological Research*, 53(3):527–51  
 Bonjour, L. (1985), *The Structure of Empirical Knowledge*, Harvard University Press, Cambridge, MA

- DeRose, K. (1996), "Knowledge, Assertion, and Lotteries," *Australasian Journal of Philosophy*, 74:568–80
- Feldman, R. (2000), "The Ethics of Belief," *Philosophy and Phenomenological Research*, 60: 667–95
- Feldman, R. (2001), "Voluntary Belief and Epistemic Evaluation," in Steup (2001b), pp. 77–92
- Feldman, R. and Conee, E. (1985), "Evidentialism," *Philosophical Studies*, 48:15–34
- Field, H. (1998), "Epistemological Nonfactualism and the A Prioricity of Logic," *Philosophical Studies*, 92:1–24
- Foley, R. (1979), "Justified Inconsistent Beliefs," *American Philosophical Quarterly*, 16:247–57
- Ginet, C. (2001), "Deciding to Believe," in Steup (2001b), pp. 63–76
- Goldman, A. (1988), "Strong and Weak Justification," *Philosophical Perspectives*, 2:51–69
- Goldman, A. (1999), "Internalism Exposed," *Journal of Philosophy*, 96(6):271–293
- Heller, M. (1999), "The Proper Role for Contextualism in an Anti-Luck Epistemology," *Philosophical Perspectives*, 13:115–30
- Kornblith, H. (2001), "Epistemic Obligation and the Possibility of Internalism," in *Virtue Epistemology*, Fairweather, A. and Zagzebski, L., eds., pp. 231–48, Oxford University Press, Oxford
- Merricks, T. (1995), "Warrant Entails Truth," *Philosophy and Phenomenological Research*, 55:841–57
- Nelkin, D. K. (2000), "The Lottery Paradox, Knowledge, and Rationality," *Philosophical Review*, 109(3):373–409
- Nozick, R. (1981), *Philosophical Explanations*, Harvard University Press, Cambridge, MA
- Plantinga, A. (1993a), *Warrant and Proper Function*, Oxford University Press, Oxford
- Plantinga, A. (1993b), *Warrant: The Current Debate*, Oxford University Press, Oxford
- Slote, M. (1979), "Assertion and Belief," in *Papers on Language and Logic*, Dancy, J., ed., Keele University Library, Keele
- Sosa, E. (1991), "Reliabilism and Intellectual Virtue," in *Knowledge In Perspective*, pp. 131–45, Cambridge University Press, Cambridge
- Sosa, E. (1999), "Skepticism and the Internal/External Divide," in *The Blackwell Guide to Epistemology*, Greco, J. and Sosa, E., eds., pp. 145–57, Blackwell, Oxford
- Steup, M. (2001a), "Introduction," in Steup (2001b), pp. 3–18
- Steup, M. (ed.) (2001b), *Knowledge, Truth, and Duty*, Oxford University Press, Oxford
- Unger, P. (1975), *Ignorance: A Case for Scepticism*, Oxford University Press, Oxford
- Williams, B. (1973), *Problems of the Self*, Cambridge University Press, Cambridge
- Williamson, T. (2000), *Knowledge and Its Limits*, Oxford University Press, Oxford
- Wright, C. (1991), "Scepticism and Dreaming: Imploding The Demon," *Mind*, 100:87–116
- Wright, C. (1996), "Response to Commentators," *Philosophy and Phenomenological Research*, 56:911–41
- Zagzebski, L. T. (1996), *Virtues of the Mind*, Cambridge University Press, Cambridge