Is Knowledge Justified True Belief?

John Turri
University of Waterloo
john.turri@gmail.com

Abstract: Is knowledge justified true belief? Most philosophers believe that the answer is clearly 'no', as demonstrated by Gettier cases. But Gettier cases don't obviously refute the traditional view that knowledge is justified true belief (JTB). There are ways of resisting Gettier cases, at least one of which is partly successful. Nevertheless, when properly understood, Gettier cases point to a flaw in JTB, though it takes some work to appreciate just what it is. The nature of the flaw helps us better understand the nature of knowledge and epistemic justification. I propose a crucial improvement to the traditional view, relying on an intuitive and independently plausible metaphysical distinction pertaining to the manifestation of intellectual powers, which supplements the traditional components of justification, truth and belief.

"The explication of knowledge as 'justified true belief', though it involves many pitfalls[,] . . . is, I believe, essentially sound."
– Wilfrid Sellars (1975: 99)

1. The End of an Era

The textbooks tell us that Edmund Gettier's 1963 paper "Is Justified True Belief Knowledge?" changed the course of epistemology by refuting the traditional view that knowledge is justified true belief (hereafter **JTB**) (e.g. Chisholm 1989: 90 ff; Moser: 1992; Feldman 2003: 25 ff.). Gettier produced two cases wherein intuitively the subject gains a justified true belief but fails thereby to know, demonstrating that justified true belief does not suffice for knowledge. Examples in this mold we call **Gettier cases**. Gettier was not the first to produce Gettier cases, but that needn't concern us here.[1]

Gettier cases follow a recipe (Zagzebski 1996: 288–9; compare Sosa 1991: 238). Start with a belief sufficiently justified to meet the justification requirement for knowledge. Then add an element of bad luck that would normally prevent the justified belief from being true. Lastly add a dose of good luck that "cancels out the bad," so the belief ends up true anyhow. It has proven difficult to explain why this "double luck" precludes knowledge.

Consider this classic Gettier case.

> (**HUSBAND**) Mary enters the house and looks into the living room. A familiar appearance greets her from her husband's chair. She thinks, "My husband is home," and

---

[1] Matilal 1986: 135–7 teaches us that the classical Indian philosopher Sriharsa constructed similar examples in the 1100s to confound his opponents. Chisholm 1989: 92–3 reminds us that Meinong and Russell produced similar examples earlier in the twentieth century.

then walks into the den. But Mary misidentified the man in the chair. It is not her husband, but his brother, whom she had no reason to think was even in the country. However, her husband was seated along the opposite wall of the living room, out of Mary's sight, dozing in a different chair. (Adapted from Zagzebski 1996: 285–6)[2]

Virtually all epistemologists intuit that Mary has a justified true belief, but does not know, that her husband is home.[3] Many regard HUSBAND and its ilk as *obvious counterexamples* to JTB.

Consider also this case.

(**BARN**) Henry and his son are driving through the country. Henry pulls over to stretch his legs, and while doing so regales his son with a list of items currently in view along the roadside. "That's a tractor. That's a combine. That's a horse. That's a silo. And that's a fine barn," Henry added, pointing to the nearby roadside barn. It was indeed a fine barn Henry saw. But unbeknownst to them the locals recently secretly replaced nearly every barn in the county with papier-mâché fake barns. Henry happens to see the one real barn in the whole county. Had he instead set eyes on any of the numerous nearby fakes, he would have falsely believed it was a barn. (Adapted from Goldman 1976: 172–3)[4]

---

[2] The case resembles Chisholm's 1989: 93 sheep-in-the-field example.
[3] Sutton 2005 denies that a Gettier subject's belief is justified.
[4] Goldman attributes the case to Carl Ginet.

Henry has a justified true belief, though intuitions divide over whether he knows.[5] But many regard BARN as a counterexample to JTB.[6]

It will be convenient to represent the argument against JTB as follows:

> (**Anti-JTB**)
> 1. If JTB is true, then the Gettier subject knows.
> 2. The Gettier subject does not know.
> 3. So JTB is not true. (From 1 and 2)

Here is the plan for the rest of the paper. Sections 2 and 3 critically evaluate recent defenses of JTB, due to Stephen Hetherington and Brian Weatherson. Section 4 presents a new defense of JTB. Section 5 explains why knowledge is nevertheless not justified true belief. Section 6 shows how very close JTB came to getting it right.

2.   Near Failure

Stephen Hetherington rejects 2 (Hetherington 1999).[7] He contends that a Gettier subject knows despite coming perilously close to not knowing, and supplements this by diagnosing intuitions to the con-

---

[5] Lycan 2006 says Henry does know, and Sosa 2007 defends a view that entails as much (see his treatment of the "kaleidoscope perceiver" and "jokester" in chapters 2 and 5).

[6] Some claim that the fake-barn case is not really a Gettier case, even though it appears to threaten JTB. I set aside this worry here.

[7] Hetherington may ultimately wish to defend only the claim that justified true belief *suffices* for knowledge, rather than JTB itself. See Hetherington 1999: 174 and 2007.

trary.

Because we are imperfect, much of our knowledge is fallible. Fallible knowing is a species of failable knowing. You **failably know** Q just in case (a) you know that Q, but (b) the following disjunction is true: either (i) your belief and your justification for Q together are consistent with Q's non-truth, or (ii) Q's truth and your justification for Q together are consistent with your not believing Q, or (iii) Q's truth and your belief together are consistent with your lacking justification for Q. Disjunct (i) corresponds to fallible knowing.[8]

Failability comes in degrees. Let a **failing world** be one where we have only two of the three aforementioned conditions for knowledge—i.e. B (*belief*) and J (*justification*) but not T (*truth*), or T and J but not B, or T and B but not J. If you actually failably know Q, then the actual world is *not* a failing world.[9] But how easily might it have turned out to be a failing world? The easier it might have, the more failable your knowledge is. In other words, the closer the closest world where you lack one of J, T or B, the more failable your knowledge.

---

[8] Reed 2002 (esp. 155 n. 21) argues that Hetherington's characterization of fallible knowing is incorrect. I set aside this worry here. Knowledge failable via (ii) or (iii) lacks a familiar name; we might call it 'diffident' or 'dogmatic' respectively, though these labels seem appropriate only when you might easily have lacked belief or justification while satisfying the other two conditions.

[9] More fully spelled out, the actual world is not a failing world relative to you, Q, and the present time. And technically worlds are too course-grained for Hetherington's purposes; *centered* worlds are more appropriate. Since nothing turns on these details I set them aside. Hetherington uses 'failure world' rather than 'failing world'.

Having thus set the stage, Hetherington's first key move is to claim, "If failable knowing can be more or less failable, then in principle there can be instances of failable knowing which are very failable" (Hetherington 1999: 571). If this is true, then the way lies open to interpret Gettier cases as instances of *very failable* knowledge, but knowledge nonetheless.

Hetherington's second key move is to diagnose philosophers' intuitions to the contrary (Hetherington 1999: 575 ff.). They sense that typical instances of knowledge *differ in quality* from what we observe in Gettier cases. They interpret this difference as that between *knowing* and *not knowing*. But here they go wrong. The difference is rather that between *failably knowing* and *very failably knowing*. The double-luck in Gettier cases ensures that the Gettier subject very easily might have lacked T despite having B and J. But by the same token the double-luck ensures that the Gettier subject actually has T, B and J. In a word, Gettier subjects don't fail to know—they very failably know (or, as I will sometimes put it, they just barely know).

I emphasize that Hetherington does not claim to have *proved* Gettier subjects know. Rather he claims to have motivated a principled alternative interpretation of Gettier cases, consistent with JTB's truth. If there are criteria for knowledge, then presumably it is possible to just barely meet them. Why not say that Gettier subjects *just barely know* rather than *fail to know*?[10]

---

[10] Compare also Hetherington 1998: 456, where he urges us to not commit the "epistemic counterfactuals fallacy" — that is, to not infer that *you actually don't know* from the fact that *you counterfactu-*

Some philosophers respond that they aren't mistaken in the way Hetherington suggests (e.g. Lycan 2006: 162). The basic intuition is just that the Gettier subject doesn't know. It does not take a detour through some indeterminate qualitative difference from a normal case of knowledge, which difference is then glossed as that between knowing and not knowing.

A different response tests whether the supposed mistake occurs in other cases of very failable knowledge, as Hetherington's diagnosis would lead us to expect. It would discredit Hetherington's diagnosis if we don't mistake *just barely knowing* for *not knowing* in these other cases.

Consider this case:

> (**CAT**) Catherine sits on her patio, contemplating the recent tragic news. She perceives a cat slinking in the yard's corner, and on that basis believes there is a cat in the yard. But the tragic news so preoccupies Catherine that she could very easily have failed to form that belief, though it would still have been true (the cat would still be there) and justified for her (she still would have perceived it).

---

ally err in a very close possible world. Neta and Rohrbaugh 2004: 401 likewise caution us against overestimating the significance of threats that "remain purely counterfactual." Of these they say, "Even though things *could* have gone epistemically less well, and almost did go epistemically less well, in point of fact, the threat was avoided and the actual case remains epistemically unproblematic." (They deny that Henry knows, however, because he faces *actual* imminent epistemic threats.)

Catherine clearly knows there is a cat in the yard. But she just barely knows (there is a very nearby failing world lacking B despite the presence of T and J). Yet this does not confuse us. We do not mistake her just barely knowing for not knowing.

We can deepen this objection by extending Hetherington's characterization of failing worlds. Recall that a failing world is one where only *two* of J, T and B are present. Let a **double-failing world** be one where only *one* of them is present. A nearby double-failing world should make for greater failability than a nearby failing world, and so make the subject come *even closer* to not knowing. If Hetherington's diagnosis is correct, then we should be even more liable to mistake these cases for not knowing.

But we do not make that mistake. Consider this case:

> (**PAIN**) The automatic door improbably malfunctions and closes prematurely, striking Dora hard on her ankle. This causes excruciating pain, on which basis Dora believes that she is in serious pain. But very easily the door could have delivered a mere glancing blow, causing only very minor discomfort rather than pain. Had it done so, due to hypochondria Dora still would have believed that she was in serious pain. (Adapted from Sosa 2007: 26.)

Dora clearly knows that she is in serious pain. But her knowledge is very failable: very easily she might have had the belief without its being true or justified. Double-failing worlds lurk nearby. Yet this does not confuse us.

Consider also this case:

(**SQUAD**) Courtney has been court-martialed and sentenced to death by firing squad. She refuses the blindfold so she can look her executioners in the eye. All ten of them aim their rifles at her. The commanding officer takes an unsteady breath and yells, "Fire!" Nearly impossibly, in unison all ten rifles *click* but fail to discharge. Courtney, who had gracefully maintained her nerve throughout, laughs aloud, "I am lucky to still be alive." (Compare the "turtle watcher" case in Unger 1968: 160 .)

Courtney clearly knows that she is still alive. But in most nearby worlds she is dead, so neither believes nor has justification that she is alive. Double-failing (indeed, *triple*-failing) worlds lurk nearby. Yet again this does not confuse us.

In light of these problems for Hetherington's diagnosis, it pays to reconsider his claim, "If failable knowing can be more or less failable, then in principle there can be instances of failable knowing which are very failable." Hetherington's opponents could concede this, but insist that *just barely knowing* is possible in some ways but not others. Gettier cases exemplify the latter sort; CAT, PAIN and SQUAD exemplify the former. Knowledge, they might say, is consistent with some but not all types of near-miss. Defenders of JTB must look elsewhere for aid in overcoming Gettier cases.

3. Counting Costs

Brian Weatherson also suggests rejecting premise 2. He concedes that intuition favors 2, but contends that broader theoretical considerations may favor rejecting it.

Weatherson proposes two main criteria for judging a philosophical theory (Weatherson 2003: 11).[11] First, to what extent does it respect our pretheoretical intuitions? Second, how simple and systematic is it? The first criterion is relatively straightforward. You check to see how often it agrees with our intuitions about actual and possible cases. You also check to see whether it possesses resources to effectively *explain away* the disagreement for the cases where it disagrees. (Just as a friend can sometimes show you respect by properly correcting you, so can a theory respect our intuitions by occasionally correcting them.) As for the second criterion, short and clear analyses are simple; analyses that illuminate a concept's relationships with other "significant" concepts are systematic.

JTB gives the intuitively wrong verdict in Gettier cases. But this is not decisive because no theory can respect all our pretheoretical intuitions (Weatherson 2003: 24). Moreover JTB might possess resources for explaining away its disagreement with intuition in these cases. And even if it possesses no such resources, it is simple and it might be significantly more simple and systematic than its rivals,

---

[11] Initially Weatherson (2003: 8–10) distinguishes four criteria. I'm counting them differently. I'm also ignoring some subtleties he mentions concerning property "naturalness," though my presentation of the second criterion seems to capture its most important aspects.

enough so to outweigh any disadvantage due to Gettier cases.

So here we have a framework for rescuing JTB. Unfortunately Weatherson stops short of fully deploying it. He discusses JTB to illustrate a point about philosophical method (i.e. counterexamples' significance), and so goes no further than claiming it is "*prima facie* plausible" that JTB excels on the second criterion (Weatherson 2003: 2, 11, 27). Nevertheless we have seen enough to judge that his strategy will not succeed.

Consider the view that knowledge is genuinely undefeated justified true belief (**GUJTB**) (Klein 1971; 1976).[12] While obviously closely related to JTB, it adds further necessary conditions. We need not go into detail about these further conditions, but suffice it to say that they cause it to return the intuitively correct verdict in standard Gettier cases. So it outperforms JTB on the first criterion. As for the second criterion, it sacrifices some simplicity to better match our intuitions; but it is *guaranteed* to be more systematic than JTB. Whatever illuminating relationships JTB reveals between knowledge and other significant concepts, GUJTB will illuminate all of those *plus* knowledge's relationship to *defeasibility*. Thus Weatherson's calculus favors GUJTB over JTB.[13]

---

[12] Sometimes Klein restricts his definition to *inferential knowledge*, but that qualification needn't concern us.

[13] A defender of JTB might object, as an anonymous referee suggested, that Weatherson's calculus does no such thing, because JTB *does* reflect knowledge's relation to defeasibility; it's just that, according to JTB, the relation is superficial. In response, however superficial the relation is, so long as there is one, JTB fails to reflect it, because it simply ignores defeasibility. So for this objection to clearly succeed, there would have to be *no* relationship between defeasib-

An analogous point is true of my positive proposal presented in sections 5 and 6 below. My proposal also adds a further necessary condition to the analysis of knowledge, which reveals its relationship to a concept fundamental to our understanding of the world, namely, that of an outcome manifesting a disposition. More on all of this later.

Thus while Weatherson's methodological observations are plausible, they offer little solace to JTB's defenders.

4. Good and Bad

A new argument defending JTB is that "bad" counterparts of Gettier subjects know Q, and these bad counterparts know Q only if Gettier subjects know Q, so Gettier subjects know Q.

Meet Bad Henry.

> (**HOOLIGAN**) Bad Henry is a hooligan who does bad things. He wants to destroy a barn. He will destroy a barn. He drives out into the country to find one. He pulls over after an hour, retrieves his bazooka, and takes aim with unerring accuracy at the roadside barn he sees. Calm, cool and collected as he pulls the trigger, he

---

ility and knowledge. But then the objection amounts to little more than simply denying that GUJTB is true. I also worry that responding this way would trivialize the criterion of being systematic, since a view could always be defended by saying, "No, my view *is* systematic because it reveals the relationship between X and Y, namely, that there is none!" In the final analysis, then, I don't think this objection succeeds.

thinks, "That sure is a nice barn . . . now was a nice barn — ha!" He destroyed the barn. He feels no remorse. He is forever after known as "Bad Henry, bane of barns." He is bad — very bad.

Bad Henry knows he is destroying a barn as he pulls the trigger. To know that, he had to know it was a barn as he took aim. So he did know it was a barn.

Now we add the twist: Bad Henry was in Fake Barn Country and just happened to shoot at the only barn around. Indeed Bad Henry destroyed the very barn that Good Henry gazed upon earlier that same day, from the very spot that Good Henry stood gazing. All the other "barns" were holograms. Nevertheless the intuition remains: Bad Henry knew he was destroying a barn. So he knew it was a barn.

It's very plausible to suppose that Bad Henry knows it's a barn only if Good Henry knows it's a barn. Bad Henry does know it's a barn. So Good Henry knows too.

Meet Bad Mary.

(**CHIP**) Bad Mary is a vindictive sadist who wants to cause her hapless husband, Benedict, excruciating pain. She *will* cause him excruciating pain. Drawing on her unsurpassed electrical engineering skills, she designs and secretly implants in Benedict's neck a small chip that administers excruciatingly painful electrical shocks. She places the device's control panel in the den, with transponders in every room in the house. The device

works only if Benedict is home, as Mary well knows.

Presently Mary enters the house and looks into the living room. A familiar appearance greets her from Benedict's chair. She thinks, "My worthless lazy husband is home — now's my chance!" and dashes into the den. Eagerly she activates the chip, which unleashes wave after wave of excruciating electrical shocks through Benedict's body. She revels in his frantic screams emanating from the living room.

She electrocuted her husband. She feels no remorse. She is forever after known as "Bad Mary, bane of Benedict." She is bad — *very* bad.

Bad Mary knew she was electrocuting her husband as she triggered that device. To know that, she had to know her husband was home. So she did know he was home.

Now we add the twist: Bad Mary misidentified the man in the chair. It isn't Benedict but his brother, whom she had no reason to think was even in the country. Benedict was dozing in a different living-room chair, out of Mary's sight. Nevertheless the intuition remains: Bad Mary knew she was electrocuting her husband. So she did know he was home.

It's very plausible to suppose that Bad Mary knows only if Good Mary (the protagonist of HUSBAND) knows. Bad Mary does know. So Good Mary knows too.

Generalizing this line of thought, we may conjecture that for any Gettier subject, there is a relevantly similar "bad" counterpart

who does know, which counterpart knows only if the Gettier subject knows. So the Gettier subject knows.[14]

Why think that the bad counterpart knows only if the Gettier subject knows? Because they have exactly the same evidence, which they use exactly the same way. If it's good enough for the one to know, it's good enough for the other.[15]

Of course, a critic might say that this cuts both ways, and so we should conclude that since it isn't good enough for the Gettier subject to know, then it isn't good enough for the bad counterpart to know either. (As the saying goes, one person's *modus ponens* is another's *modus tollens*.) But I doubt that JTB's critics would pretend that it's as obvious to them that we should apply *modus tollens* here, as it is that the original Gettier subjects don't know. In other words, we've shifted to terrain much more favorable to JTB. JTB's lonely defenders will certainly view this as progress.

One response to the cases is to liken the claim that Bad Mary knew she was electrocuting her husband (or the claim that Bad

---

[14] Beebe and Buckwalter (forthcoming) report experimental results suggesting that people "are more likely to say that agents know their actions will bring about certain side-effects, if the effects are bad than if they are good," which is analogous to Joshua Knobe's (2003) interesting finding that people are more likely to say that an agent intentionally brought about a bad side-effect than a good one. Beebe and Buckwalter's work supports the claim that Bad Henry knows he is destroying a barn, and that Bad Mary knows she is electrocuting her husband.

[15] I imagine that those attracted to the view that your "practical environment" can affect what you know might have principled grounds for disagreeing with the claim expressed by this conditional. See e.g. Fantl and McGrath 2002, 2007; Hawthorne 2004: Chapter 4; and Stanley 2005: Chapter 5.

Henry knew he was destroying a barn) to the claim that Medieval European peasants knew that the Earth was flat.[16] The latter claim is either false or involves a sense of 'know' different from the sense epistemologists are interested in. But this maneuver doesn't succeed. On the one hand, if it's false that Medieval peasants knew the Earth was flat, then that's because it was *false* that the Earth was flat, and you can't know false things. But Gettiered beliefs are by definition *true*, so it's unclear why we should treat the two cases similarly. On the other hand, if it's a true claim involving a different and epistemologically irrelevant sense of 'know', then that sense of 'know' is non-factive (otherwise, the claim that the peasants knew would be false). But again, Gettiered beliefs are by definition *true*, so it's unclear why we would employ the non-factive sense to describe the Gettier subject. Much more would need to be said in order to make the analogy plausible.

Some people have suggested to me that in the context of evaluating action, 'know' might be ambiguous between the ascription of a *cognitive achievement* on the one hand, and *intentional action* on the other.[17] When we say that Bad Henry knew he was blowing up a barn, we mean that he *intentionally blew up a barn*, but this doesn't entail that he knew it was a barn. This response will seem *ad hoc* unless it's backed up by actual linguistic evidence of ambiguity. The response also strikes me as implausible, since this conjunction sounds terrible: 'Henry knows he's destroying a barn, but he

---

[16] Thanks to Allan Hazlett for conversation on this point.

[17] As E.J. Coffman and Sharifa Mohamed independently suggested to me.

doesn't know it's a barn'. The most obvious explanation for why it sounds terrible is that it's contradictory. *Henry's knowing that he's destroying a barn* straightforwardly entails *his knowing that it's a barn*. Perhaps there is some other explanation of the conjunction's infelicity, but it's the critic's burden to produce it.

A better response to the argument would be to say that the moral dimension of the "bad" cases causes a performance error by polluting our intuitions about whether the bad people know. We recognize that these bad people deserve strong censure for their actions, and one way of accentuating the blame they deserve is to say that they "knew" that they were about to do something bad. It makes it sound worse to say that someone *knew* his action would have a certain awful consequence, than if you say that he was justified in thinking that his action would have that awful consequence. The knowledge-ascription is literally false, but serves a purpose.

This last response strikes me as having some merit. But if it's the best response available, then JTB's proponents will consider themselves to have made real progress. They will have gone from confronting what was supposed to be an irresistible counterexample to JTB, to confronting a somewhat plausible, but by no means clearly successful, attempt to undermine intuitions about the bad counterparts.

Even though I agree that JTB is subtly wrong (in a way I'll explain momentarily), I still think that the argument reviewed in this section ought to make us reconsider the intuition that the Gettier subject doesn't know. We should be less confident now than we

were before that the Gettier subject doesn't know. To that extent, the argument succeeds. And given how widely and unquestioningly accepted the intuition is, this is a noteworthy development.

5.   Why Knowledge Is Not Merely Justified True Belief

In these final two sections, I would like to suggest a way of evaluating JTB once we've given up on the claim that it suffers irredeemably from counterexamples. There are always ways to resist purported counterexamples. Although I do think JTB is on the right track, I also think it's false because it omits something important about the nature of knowledge. The fact that it overlooks something essential to knowledge naturally leads us to suspect that it will face genuine counterexamples. But counterexamples are diagnostic tools. They can suggest that an analysis has gone wrong, without necessarily telling us where it goes wrong. An account of the underlying problem is more valuable than an example suggesting that such a problem exists. I'll now try to provide such an account.

Start with the observation that gaining knowledge involves the exercise of intellectual abilities or powers (compare e.g. Aristotle's *Posterior Analytics*; Reid 1764, 1785; Kant 1781; Sellars 1956). We gain perceptual knowledge in virtue of exercising our perceptual abilities. We gain knowledge of necessary truths, and knowledgeably deduce consequences therefrom, in virtue of exercising our rational powers. We gain introspective knowledge of our own mental condition in virtue of exercising our reflective powers. As far as I

can tell, there is no instance where knowledge is gained without the subject exercising her intellectual powers.[18]

Abilities and powers are dispositions.[19] For an outcome to manifest a disposition, it isn't enough for the outcome to be due to the disposition in just *any* way. We have fairly consistent intuitions about this across a broad range of cases. Consider this pair of cases:

> (**BROKE**) I sat at the table feeding baby Mario his breakfast. Unfortunately my glass of orange juice was within reach, and Mario swatted it off the table. Spoon in one hand, baby in the other, I helplessly watched the glass tumble down, down, down. It *broke*.

> (**PACKED**) My wife and I were excited as we packed our belongings to move into our new home. "Remember, we must carefully pack this glass, because it is very fragile," said my wife, holding up her favorite glass. I followed her advice. The glass was *carefully packed*.

In each case the outcome is in some way due to the glass's fragility. (Neither outcome obtains *only* because of fragility — in BROKE Mario and the floor help out, in PACKED my effort also contributes

---

[18] Would innate knowledge be a counterexample to this claim? Not really, because it doesn't seem like we *gain* innate knowledge. Rather, we're created with it. And we then *retain* it via our power of memory.

[19] Or if they're not types of disposition, they still share with dispositions the crucial feature I'm concerned with below, namely, that certain outcomes manifest them. This is all that matters for present purposes, so for simplicity in the main text I will continue mainly speaking of dispositions.

— but that doesn't spoil the point.) Yet we all recognize an important difference: the outcomes are not due *in the same way* to fragility. In BROKE the glass broke because it was fragile, and its breaking manifests its fragility. In PACKED the glass was carefully packed because it was fragile, but its being carefully packed does not manifest its fragility. *Breaking* is the right sort of outcome to manifest fragility, but *being carefully packed* isn't.

Consider also this pair of cases.

(**BOIL**) You place a cup of water in the microwave and press start. The magnetron generates microwaves that travel into the central compartment, penetrate the water and excite its molecules. Soon the water *boils*.

(**FIRE**) You place a cup of water in the microwave and press start. The magnetron generates microwaves that cause an insufficiently insulated wire in the control circuit to catch fire, which fire deactivates the magnetron and spreads to the central compartment. Soon the water *boils*.

Both outcomes are in some way due to the microwave's boiling power. But again, we all recognize an important difference. The outcome in BOIL manifests the microwave's boiling power, whereas the outcome in FIRE does not. We have a plain way to mark the distinction in ordinary language: in the former case, but not the latter, the microwave *boils the water*.

The pairs of examples highlight a general distinction between

(A) an outcome happening *merely because* of a disposition and (B) an outcome *manifesting* a disposition. No metaphysical theory teaches us this distinction. We understand it perfectly well pretheoretically.

Since the distinction between A and B is a perfectly general one, it applies to intellectual powers and abilities too. One central outcome of the exercise of intellectual powers is the formation of true beliefs. We can ask how such an outcome — namely, the formation of a true belief — relates to the relevant intellectual power. For instance, we can ask whether it manifests the relevant power. If it fails to do so, then we wouldn't expect the subject to gain knowledge, just as in FIRE the microwave doesn't boil the water.

Notice a striking similarity between the "double luck" recipe for generating Gettier cases and what happens in FIRE. FIRE exemplifies that *exact same* pattern. The microwave initiates a process that would normally result in the water's boiling. Bad luck strikes: the magnetron is disabled, which would normally result in the water not boiling. But then "good" luck strikes: the damaged circuit starts afire, resulting in the water's boiling anyhow. Double luck prevents the outcome from manifesting the relevant power. Exactly the same thing happens in Gettier cases. This should increase our confidence that we've correctly identified a more general pattern of failure that Gettier cases belong to.[20]

Given that JTB ignores the central role of intellectual powers in the production of knowledge, it's no wonder that the most difficult

---

[20] For more discussion, see Turri forthcoming.

cases facing JTB involve a defective relationship between a subject's intellectual powers and the true belief that their operation helps produce. Of course, there are clever ways of trying to offset the intuitive cost JTB incurs as a result of Gettier cases, several of which we considered earlier. But none of those strategies can offset the fact that JTB misses something fundamentally important about the nature of knowledge, namely, the particular way that intellectual powers must relate to true beliefs, and how that fits into a more general pattern relating dispositions to outcomes.

Notice that even though my diagnosis of JTB's omission is motivated on grounds independent of Gettier's discussion, it nevertheless can explain why Gettier cases pose trouble for JTB. Indeed, equipped with the view that knowledge is true belief manifesting intellectual power, one could predict that JTB will give the wrong verdict precisely in Gettier-style cases. Combine this with the fact that the proposal emerges naturally from a fairly innocent observation shared among a diverse group of distinguished philosophers, and it becomes clear that the proposal merits serious consideration.

6. Explaining JTB's Persistent Appeal

Wilfrid Sellars once remarked that despite the Gettier problem, he still thought JTB was "essentially sound." The traditionalists, Sellars thought, were basically right when they said, "K = JTB."

If we make but one assumption, then my proposal would vindicate both Sellars and the tradition by showing JTB to be *almost*

right. Traditionalists had all the right pieces in place, but failed to fully appreciate how they must be interrelated to yield knowledge.

Assume that a person's intellectual powers or abilities are the source of epistemic justification for her.[21] Given that assumption, to say 'knowledge is true belief manifesting intellectual power' is basically to say 'K = J⇨TB', where the arrow represents the relation of manifestation. This addition constitutes a small but crucial improvement on the traditional view, relying only on an intuitive and independently plausible metaphysical distinction to supplement the traditional components.

I think this speaks greatly in favor of the assumption that intellectual powers are the source of epistemic justification. That is, the fact that the assumption reveals JTB to be almost right counts in its favor. There's a reason why so many smart people found JTB so attractive for so long. Once we're convinced that JTB is false, the next most plausible explanation for the attraction is that JTB is *close* to being true. And if my analysis here is correct, then JTB hit very close to the mark indeed.[22]

---

[21] Compare Sosa 1991, Greco 1993, Zagzebski 1996, and Bloomfield 2000.

[22] For helpful conversation and feedback on relevant material, I am happy to thank Margaux Carter, E.J. Coffman, John Greco, Allan Hazlett, Stephen Hetherington, Sharifa Mohamed, Ram Neta, Duncan Pritchard, Andrew Rotondo, Ernest Sosa, Angelo Turri, and an anonymous referee.

References

Aristotle. *Posterior Analytics*. Trans. G.R.G. Mure. In *The Basic Works of Aristotle*. Ed. Richard McKeon. New York: Random House, 1941.

Beebe, James and Wesley Buckwalter. "The Epistemic Side-Effect Effect." Forthcoming in *Mind and Language*. <http://www.acsu.buffalo.edu/~jbeebe2/Beebe%20Buckwalter%20ESEE.pdf>. Accessed 23 July 2010.

Bloomfield, Paul. 2000. "Virtue Epistemology and the Epistemology of Virtue." *Philosophy and Phenomenological Research* 60.1: 23–43.

Chisholm, Roderick M. 1989. *Theory of Knowledge*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall.

Fantl, Jeremy and Matthew McGrath. 2002. "Evidence, Pragmatics, and Justification." *The Philosophical Review* 111.1: 67–94.

Fantl, Jeremy and Matthew McGrath. 2007. "On Pragmatic Encroachment in Epistemology." *Philosophy and Phenomenological Research* 75.3: 558–589.

Feldman, Richard. 2003. *Epistemology*. Upper Saddle River, NJ: Prentice Hall.

Gettier, Edmund. 1963. "Is Justified True Belief Knowledge?" *Analysis* 23.6: 121–3.

Goldman, Alvin. 1976. "Discrimination and Perceptual Knowledge." *The Journal of Philosophy* 73.20: 771–791.

Greco, John. 1993. "Virtues and Vices of Virtue Epistemology." *Canadian Journal of Philosophy* 23: 413–32.

Knobe, Joshua. 2003. "Intentional Action and Side Effects in Ordinary Language." *Analysis* 63.3: 190–4.

Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.

Hetherington, Stephen. 1998. "Actually Knowing." *The Philosophical Quarterly* 48.193: 453–469.

Hetherington, Stephen. 1999. "Knowing Failably." *The Journal of Philosophy* 96.11: 565–587.

Hetherington, Stephen, ed. 2006. *Epistemology Futures*. Oxford: Oxford University Press.

Hetherington, Stephen. 2007. "Is This a World Where Knowledge Has to Include Justification?" *Philosophy and Phenomenological Research* 75.1: 41–69.

Kant, Immanuel. 1781 [1996]. *Critique of Pure Reason*. Trans. Werner S. Pluhar. Indianapolis: Hackett.

Klein, Peter. 1971. "A Proposed Definition of Propositional Knowledge." *The Journal of Philosophy* 68.16: 471–482.

Klein, Peter. 1976. "Knowledge, Causality, and Defeasibility." *The Journal of Philosophy* 73.20: 792–812.

Lycan, William G. 2006. "On the Gettier Problem problem." In Hetherington, ed.

Neta, Ram and Guy Rohrbaugh. 2004. "Luminosity and the Safety of Knowledge." *Pacific Philosophical Quarterly* 85.4: 396–406.

Matilal, Bimal Krishna. 1986. *Perception: An Essay on Classical Indian Theories of Knowledge*. Oxford: Oxford University Press.

Moser, Paul K. 1992. "The Gettier Problem." In *A Companion to Epistemology*, ed. Jonathan and Ernest Sosa. Malden, Mass: Blackwell.

Reed, Baron. 2002. "How to Think about Fallibilism." *Philosophical Studies* 107.2: 143–157.

Reid, Thomas. 1764 [1997]. *An Inquiry into the Human Mind on the Principles of Common Sense*. Ed. Derek Brookes. University Park, Penn.: Pennsylvania State University Press.

Reid, Thomas. 1785 [2002]. *Essays on the Intellectual Powers of Man*. Ed. Derek Brookes. Edinburgh: Edinburgh University Press.

Sellars, Wilfrid. 1956. "Empiricism and the Philosophy of Mind." In

Science, Perception and Reality, Atascadero, Calif.: Ridgeview, 1963.

Sellars, Wilfrid. 1975. "Epistemic Principles." In *Action, Knowledge, and Reality* (Indianapolis: Bobbs-Merrill), ed. H. Castañeda. Reprinted in *Epistemology: An Anthology* (Blackwell, 2008), ed. Ernest Sosa, Jaegwon Kim, Jeremy Fantl, and Matthew McGrath.

Sosa, Ernest. 1991. *Knowledge in Perspective*. Cambridge: Cambridge University Press.

Sosa, Ernest. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge*, v. 1. Oxford: Oxford University Press.

Stanley, Jason. 2005. *Knowledge and Practical Interests*. Oxford: Oxford University Press.

Sutton, Jonathan. 2005. "Stick to What You Know." *Noûs* 39.3: 359–396.

Turri, John. Forthcoming. "Manifest Failure: The Gettier Problem Solved." Forthcoming in *Philosophers' Imprint*.

Unger, Peter. 1968. "An Analysis of Factual Knowledge." *The Journal of Philosophy* 65.6: 157–70.

Unger, Peter. 1975. *Ignorance: A Case for Skepticism*. Oxford: Oxford University Press.

Weatherson, Brian. 2003. "What Good Are Counterexamples?" *Philosophical Studies* 11.1: 1–31.

Zagzebski, Linda. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.