Generics, Epistemic Luck, and Knowledge

by

Joe Milburn

B.A. in Philosophy and Spanish, Duquesne University, 2002

M.A. in Philosophy and Linguistics, King's College London, 2006

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

PhD in Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Joe Milburn

It was defended on

May 4, 2015

and approved by

Edouard Machery, Professor of History and Philosophy of Science, The University of Pittsburgh

John McDowell, Distinguished University Professor of Philosophy, The University of Pittsburgh

Dissertation Directors: Karl Schafer, Associate Professor of Philosophy, The University of Pittsburgh

and

Kieran Setiya, Professor of Philosophy, MIT

GENERICS, EPISTEMIC LUCK, AND KNOWLEDGE

Joe Milburn, PhD

University of Pittsburgh

Philosophers generally agree that an individual cannot have knowledge that p if her belief that p is only accidentally true. A common way of putting this claim is that knowledge is incompatible with epistemic luck. After criticizing modal accounts of epistemic luck, Ioffer a novel account of knowledge precluding epistemic luck. My account appeals to *generic* facts concerning how human beings arrive at the truth. Given that it is a generic fact that human beings arrive at true beliefs through perception, memory, testimony, and inference (abductive, inductive, and deductive), an individual S's belief that p is subject to epistemic luck just in case she arrives at a true belief, but fails to *exemplify* one of these generic facts concerning human beings. Apart from offering an account of epistemic luck that is extensionally adequate and that could be put to use in a reductive account of knowledge, my account of epistemic luck reorients epistemology away from a disembodied Cartesian rationalism towards a moderate naturalism.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am much indebted to Profs. Karl Schafer and Kieran Setiya and the other members of my committee for four years of advice and support. This includes Profs. James Allen, Andrew Chignell, and Duncan Pritchard, who do to university policy and scheduling difficulties could not be on my final committee. I wish to give special thanks to Kieran Setiya, without whose encouragement graduate school would have been impossible.

1.0 CHAPTER OVERVIEW

The following dissertation consists of five chapters, all of which were written to stand alone as independent articles. But while these chapters can be read independent of each other, they all bear on either of the same two themes. The first theme is epistemic luck, i.e., the kind of luck that is paradigmatically present in Gettier cases, and which philosophers have generally held to be incompatible with knowing. The second theme is the importance of generic propositions for the work of epistemology. In particular I argue that by appealing to generic propositions concerning human beings, we can solve two long standing problems in contemporary epistemology, namely the Gettier problem and the generality problem for reliabilism.

These themes (epistemic luck and the importance of generic propositions for epistemology) I have inherited from Duncan Pritchard and Kieran Setiya respectively. Pritchard's work on epistemic luck is well known, and I will not speak more of it here, as it will be treated at length in the second chapter 'Safety, Gettier Cases, and Epistemic Luck'. I *will* speak more of Setiya's work before giving a brief summary of the chapters, however, because, while Setiya's work was the immediate catalyst for both of my chapters 'Truth Connectedness, Generics, and the Gettier Problem,' and 'Generics and the Generality Problem,' I do not directly engage with Setiya in either chapter.

To my knowledge, Setiya is the only author in the contemporary literature that has considered the importance of generic propositions to epistemology. Setiya appeals to generic

propositions in dealing with a skeptical argument directed at our ethical knowledge. The skeptical argument goes as follows:

1. We can have ethical knowledge only if our ethical beliefs are non-accidentally true.

2. But our ethical beliefs are non-accidentally true only if we form them using a reliable method of belief formation, and there is an explanatory connection between the reliability of the method of belief formation and our using it, i.e., either we use the method *because* it is reliable, or the method is reliable *because* we use it.

3. But there is not an explanatory connection between the reliability of our methods of ethical belief formation and our using these methods if the facts of ethics are irreducible, causally inert, and constitutively independent of us.

4. Therefore, our ethical beliefs are at best accidentally true and we do not have ethical knowledge. (For more details of this argument see Setiya 2012 ch. 3.)

Setiya argues that we can defeat this skeptical argument while allowing for the possibility of widespread and persistent ethical disagreement by appealing to what can be called *Natural Externalism*. According to Natural Externalism part of what it is to have ethical concepts is to belong to a life-form whose methods for forming ethical beliefs are sufficiently reliable. Given Natural Externalism we can give the following explanation for why we use the methods we do that appeals to their reliability: We use method m for forming ethical beliefs, because we are human beings, and human beings use method m for forming ethical beliefs. But if m were not reliable, then we would not use it to form ethical beliefs, since if it were not reliable, human beings would no longer have the concept of morality. So the skeptical challenge given above can be met. At the same time, this thesis allows for widespread disagreement between members of the human life form regarding ethical beliefs. This is because generic statements like 'Human beings use method m for forming ethical beliefs' can admit of counter-instances but still be true.

I must admit that I find the thesis of Natural Externalism somewhat *ad hoc*. While it does allow us to defeat the skeptical argument concerning ethical knowledge while allowing for the possibility of widespread disagreement between individuals and communities, as a thesis concerning conceptual capacities, I find it odd. If widespread and radical disagreement regarding ethics can occur between individual human beings, or whole communities of human beings, it seems it also could occur between different life forms. But Natural Externalism denies that this is a possibility. Furthermore, consider that there could be an individual that belonged to a life-form that did not have a reliable method for forming ethical beliefs; however, while the individual's life-form did not use a reliable method in regard to forming ethical beliefs, we might have overwhelming evidence that the individual used a reliable method. When asked whether or not e.g. one should ever intentionally kill an innocent human being, this individual would answer correctly, and he would always act in accordance with his professed ethical beliefs. But according to Natural Externalism, this individual could have no ethical concepts and no ethical beliefs, and so would not have a reliable method for forming ethical beliefs. This too strikes me as counterintuitive. Finally, it certainly is not the case that anything like Natural Externalism holds true in general. Oftentimes those who make up the scientific elite, for example, possess concepts such that it simply is not the case that the following generic is true: *human beings use reliable methods to form beliefs involving these sorts of concepts*.

That being said, Setiya's claims about the conditions on knowledge seem to me to be fully credible. Of course I think knowledge must be non-accidentally true, or in other words, not subject to epistemic luck. But at the same time, I find the idea that there must be an explanatory connection between an individual S using a method m, and the reliability of method m in order for S's belief that p formed by method m to be non-accidentally true, compelling. As will

become clear in the chapter 'Generics and the Generality Problem' I believe that a method of belief formation is *ipso facto* reliable if there is a true generic statement *human beings use method m to form true beliefs.* It follows that my account of epistemic luck is in line with Setiya's core idea regarding accidentally true belief. Of course, what I have taken most from Setiya's epistemological work is the idea that generic propositions have an important role to play in epistemology. Of this I am convinced, though others must judge if what follows bears this out.

Here I briefly summarize the chapters of my dissertation. In the first chapter, 'Subject-Involving Luck', I resist Nathan Ballantyne's recent arguments that a better understanding of luck will *not* afford us a better understanding of knowledge. Ballantyne's argument, in brief, is that there is no reason for supposing that it must be lucky for an individual S to believe the truth in Gettier cases. This is because whether or not it is lucky for S that she believes the truth that p will depend on whether it is good for S to believe the truth that p. But it could be that an individual S's belief that p is Gettierized, while it is not good for S to believe truly that p. In this case, it is not lucky for S that she believes truly that p even though her belief is Gettierized. My response to Ballantyne is to shift focus from accounts of what I call subject-relative luck, i.e., accounts of what it is for an event to be lucky for an individual S, to accounts of subject-involving luck, i.e., accounts of what it is for it to be a matter of luck that an individual S Φs. Once we make this shift, there is no reason for doubting that all Gettierized beliefs are subject to epistemic luck. That is, we have no reason to doubt that it is a matter of luck that Gettierized beliefs are true. Furthermore, I argue that subject-involving luck is what is at stake in so-called cases of moral luck.

In the second chapter, ‘Safety, Gettier cases, and Anti-Luck Epistemology’ I give a novel argument against safety accounts of epistemic luck. According to safety accounts of epistemic luck, an individual S’s belief that p is subject to epistemic luck just in case her belief is unsafe, i.e., just in case in forming her belief that p as she did, S was at (serious) risk of forming a false belief. Recently, a number of counterexamples against safety accounts of epistemic luck have been introduced in the literature. These putative counterexamples are supposed to be instances of safe Gettierized belief. In this chapter I argue that these putative counterexamples fail, since they rely on questionable assumptions about the methods of belief formation that an individual uses in forming her Gettierized belief. At the same time, however, I argue that we have principled reason for believing that there can be safe Gettierized belief. My argument relies on the following two premises. The first premise is that we can specify Gettier cases without making either an explicit or implicit appeal to unsafe belief. The second premise is a weakened form of Hume's dictum (WHD). According to (WHD) if two types of contingent events (or states of affairs) are wholly distinct then we have (ceteris paribus) sufficient reason for denying there is a necessary connection between these types of events. It follows from these premises that we have sufficient reason for believing that there can be safe Gettierized beliefs.

The third paper ‘Against Mixed Epistemology’ takes on some hybrid reductive views of knowledge that have appeared lately in the literature. These views understand an individual S’s belief that p to be an instance of knowledge just in case S’s belief that p meets a safety condition and if S’s belief that p meets an independent ability condition on knowledge. I argue that this kid of mixed account of knowledge is untenable. The putative counterexamples to a pure safety account of knowledge are equally problematic for a mixed account of knowledge that includes an ability condition, and vice versa. In this chapter I tentatively endorse the possibility that a broad

kind of reliabilist contextualism can deal with the problems facing both pure safety and pure ability accounts of knowledge, while being open to the idea that perhaps pure safety-based, pure ability-based, and mixed epistemologies fail to give an adequate account of knowledge. The sequel shows my considered view is that we can give something like a pure ability-based account of knowledge without making recourse to contextualism.

The fourth paper 'Truth-Connectedness, Generics, and the Gettier Problem' is the central piece of the dissertation where the two guiding themes -- epistemic luck, and the importance of generics for epistemology – intersect. After noting that most philosophers take subjects in Gettier cases to lack knowledge because the grounds for their beliefs are not properly related to the truth of their beliefs, and after reviewing some of the difficulties for standard modal and causal/explanatory accounts of proper truth-connectedness, I provide an account of proper truth-connectedness in terms of generic propositions. In so doing, I articulate and defend a novel account of luck that I call the generic account of luck, which is an account of subject-involving luck. According to the generic account, it is a matter of luck that S Φs (or alternatively, it is an accident that S Φs) just in case S is an F, and there is a true generic proposition representing a principled connection between Fs and Φing but S does not exemplify this true proposition. In turn, an individual S's true belief that p has proper truth-connectedness just in case S is an F, there is a true generic proposition representing a principled connection between Fs and believing the truth, and in believing truly that p S exemplifies this generic proposition. I argue that this account of proper truth-connectedness offers us an extensionally adequate solution to the Gettier problem and allows for an adequate explanation of why Gettierized beliefs are not knowledge.

In the fifth and final chapter, "Generics and the Generality Problem" I again appeal to generic propositions, this time to take on the generality problem. There I claim that an individual

S's belief that p is justified just in case it is formed by a relevant, reliable, method of belief formation in the right way. A method of belief formation is relevant, if there is a true generic proposition that human beings use this method of belief formation to form beliefs. I argue that a method of belief formation is reliable just in case there is a true generic proposition that human beings use this method to form true beliefs. Finally, I argue that an individual forms a belief in the right way just in case she exemplifies the generic proposition that human beings use the relevant method of belief formation.

One of the features of these last two chapters that I am pleased with is that they emphasize that we, you and I, are human animals, and that in order to understand knowledge and justified belief we must make recourse to generic facts concerning the human animal. I am happy to do epistemology under a banner that reads: *theories of human knowledge for human beings!* No doubt, this aspect of my work will draw protests and set it apart from the mainstream of epistemic theorizing, where even those epistemologists most dedicated to philosophical naturalism give accounts of justification and knowledge that would seem to work as well for angels and rational wolves, as for human beings.

In concluding this introduction I should note that, as each chapter is designed to stand alone, there are some redundancies. I have tried to keep these to a minimum.

2.0 CHAPTER 2 SUBJECT-INVOLVING LUCK

2.1 INTRODUCTION

In recent years, philosophers have tended to think of luck as being a relation between an event (taken in the broadest sense of the term) and a subject.[1] According to these philosophers, to give an account of luck is to fill in the right hand side of the following biconditional: an event *e* is lucky for a subject S if and only if ____. We can call such accounts of luck subject-relative accounts of luck, since they attempt to spell out what it is for an event to be lucky relative to a subject.

In this essay, I want to argue that we should understand subject-relative luck as a secondary phenomenon. What is of philosophical interest is giving an account of *subject-involving luck,* i.e., filling in the right hand side of this biconditional: it is a matter of luck that S φs iff ____.

I take it that this view is contrary to nearly all philosophers who are currently interested in the topic of luck. Most of these philosophers are, to use Nathan Ballantyne's word, *interventionists* (Ballantyne 2014). They believe that by giving an account of luck in general, we will be able to intervene in a number of philosophical issues. Ballantyne, on the other hand, has recently pushed an anti-interventionist line (Ballantyne 2011, 2014). In particular, he has forcefully argued that understanding luck will not help us understand knowledge or Gettier

[1] See, for example, Statman (1991), Rescher (1995), Latus (2003), Pritchard (2005), Coffman (2007), Riggs (2009), Steglich-Petersen (2010), Levy (2011), and Ballantyne (2011).

problems. As a result, the topic of luck fails to have the broad philosophical significance many philosophers believe it does.

I agree with the majority of philosophers working on the topic of luck today (and disagree with Ballantyne) that luck is an important topic, one capable of illuminating a wide range of philosophical issues, especially those in epistemology. But I agree with Ballantyne (and disagree with the majority) that giving an account of what it is for an event *e* to be lucky for an agent S is philosophically unimportant. The problem, I argue, is that most philosophers have overlooked the difference between subject-relative and subject-involving luck. It is this latter phenomenon that is philosophically important.

If my argument works, then we should shift focus from giving an account of subject-relative luck to giving an account of subject-involving luck. One of the upshots of refocusing our energies is that *lack of control accounts of luck* (LCALs) become much more attractive. In particular, when we focus on giving an account of subject-involving luck, a wide range of classic counterexamples to lack of control accounts of luck fall by the wayside.

My essay proceeds as follows. In section 2, I explain the difference between subject-relative luck and subject-involving luck. In section 3, I argue that we should understand epistemic and moral luck as instances of subject-involving luck, not subject-relative luck. The upshot of this argument is that philosophers should be primarily interested in giving an account of subject-involving luck. In section 4, I argue that, while LCALs for subject-relative luck fall prey to counterexamples involving regularly occurring events that are beyond our control, LCALs for subject-involving luck do not face this problem. Finally, in section 5, I consider two objections to my argument and respond to them.

2.2 SUBJECT-RELATIVE LUCK VERSUS SUBJECT-INVOLVING LUCK

Consider the following expressions:

> It was lucky for Seth McFarlane that he missed his flight on September 11th.
>
> It was unlucky for Christopher Walken that he turned down the role of Han Solo.
>
> It was unlucky for Philip II, but lucky for Queen Elizabeth's subjects, that a storm destroyed the Spanish Armada.[2]

Such expressions ascribe what can be called subject-relative luck. Accounts of luck in the current literature tend to be accounts of subject-relative luck. According to these accounts, luck is a property of an event in relation to an individual. These accounts of luck then seek to elucidate this property by filling in the right-hand side of the following biconditional:

> An event *e* is lucky for a subject X iff ____.

It should be noted that expressions attributing subject-relative luck are not the only "luck" expressions in common use. Other common expressions include the following:

> Seth McFarlane was lucky that he missed his flight on September 11th.
> (ascription of luckiness)
>
> It is a matter of luck that Seth McFarlane missed his flight on September 11th.
> (ascription of subject-involving luck)

It is plausible that expressions that ascribe luckiness are not importantly different from expressions that ascribe subject-relative luck. In the one, luck is understood as a property of an event in relation to a subject; in the other, luck is understood as a property of a subject in relation to an event. The following biconditional seems to hold:

[2] This last example comes from Rescher (1995, 20).

E is lucky for S iff S is lucky that E.

In this case, if it was lucky for Seth McFarlane that he missed his flight on September 11th, then it follows that McFarlane was lucky that he missed his flight on September 11th, and vice versa. The difference between asserting one or the other of these statements is insignificant.

But this is not true of expressions that ascribe what I am calling subject-involving luck; it seems that these are importantly different from expressions that ascribe subject-relative luck or subject luck. First, on a natural way of understanding things, subject-involving luck is not a property of an event in relation to a subject, or a property of a subject in relation to an event. Rather it is a property of an event that involves a subject. What is lucky is simply this: That S φs (for example, that Seth McFarlane misses his flight). Because of this it could be a matter of luck that Seth McFarlane missed his flight on September 11th, while it was neither lucky nor unlucky *for* McFarlane that he missed his flight.

This point becomes clearer when we consider what Ballantyne (2011, 2012) calls the "significance condition" on subject-relative luck. Ballantyne points to unanimous agreement among philosophers working on luck that for an event *e* to be lucky for S, *e* must somehow be significant for S. Exactly what it is for an event to be significant to a subject S is open to debate.[3] However, that there is some significance condition for subject-relative luck is highly intuitive.

It is important to note that there is no significance condition for subject-involving luck. Perhaps missing his flight is a matter of complete indifference for Seth McFarlane; his subjective and objective well-being remains the same whether or not he misses the flight. Still, it could be

[3] See Ballantyne (2012) for some of the moves in this debate.

asked why McFarlane missed his flight, and it could be truly said in response that it was merely a matter of luck that McFarlane did so.

Or consider the following example. Suppose that the artist in residence at your university conducts an Absurdist Raffle as a work of performance art. He assigns every student and faculty member in the university a number, puts these numbers in a very large hat, and draws one. To "reward" the winner, the artist gives a member of the university administration $100,000 in Monopoly money.

Suppose that you are the winner of the Aburdist Raffle, but that you never find out about it. Surely it is neither lucky nor unlucky for you that you won the raffle. After all, we can suppose that winning the Absurdist Raffle is and should be completely insignificant to you. But while it is neither lucky nor unlucky for you that you won the raffle, it does seem to be a matter of luck that *you* won the raffle. Indeed, it is a paradigmatic example of something that happens as a matter of luck. So for subject-involving luck, unlike subject-relative luck, there is no significance condition.

2.3 EPISTEMIC LUCK AND MORAL LUCK ARE INSTANCES OF SUBJECT-INVOLVING LUCK

In the last section I made a distinction between subject-relative luck and subject-involving luck. This distinction is important to note when one considers why philosophers are currently interested in luck.

Take for instance Duncan Pritchard's anti-luck epistemology (Pritchard 2005, 2007). Pritchard's program can be understood as starting from a basic platitude, namely, "knowledge excludes luck," and building up from this platitude an account of knowledge. Doing this, says Pritchard, involves three steps: "First, we offer an account of luck. Second, we specify the sense in which knowledge is incompatible with luck. Finally, third, we put all this together to offer an anti-luck analysis of knowledge" (2007, 278). So Pritchard's account of luck is given with the ultimate aim of providing an account of knowledge.

Or consider Neil Levy's (2011) sustained argument against our being morally responsible agents. Again, Levy's program is to start with a basic platitude, that luck excludes moral responsibility, and to build up from this platitude some substantial philosophical thesis, namely, that we are never morally responsible. So Levy also has ulterior motives, so to speak, in giving a general account of luck. And by and large these cases generalize. I am aware of no contemporary philosopher who gives an account of luck who is not motivated directly or indirectly by issues involving epistemic or moral luck.

One question that is worth asking, then, is this: Is it subject-involving or subject-relative luck that is relevant for debates concerning moral or epistemic luck? I think the answer is clear: it is subject-involving luck. It follows that philosophers should be primarily concerned with giving an account of subject-involving luck.

Let's start with epistemic luck. Consider a paradigmatic example of epistemic luck: Gettierized belief.

Ballantyne (2011) shows us that understanding epistemic luck as an instance of subject-relative luck creates a number of difficulties. Consider the following example:

> Green and Blue are . . . looking over the field, gazing at what appears to be a sheep. They form true beliefs that a sheep is in the field. Though a sheep is just out of view, what Green and Blue see is a dog masquerading as a sheep. Suppose that the very same epistemic facts hold for Green and Blue, but that different significance facts hold for them. Having a true belief regarding that sheep is good for Green. . . . Not so for Blue. He is apathetic and . . . deeply depressed. In fact, suppose that Blue has no preference or desire to have a true belief that sheep is in the field; he just doesn't care. So, having a true belief about the sheep is neither good nor bad for him (Ballantyne 2011, 493–94).

If we allow that it is insignificant for Blue that he believes truly that there is a sheep in the field, then if we understand epistemic luck as an instance of subject-relative luck, it follows that Blue's Gettierized belief is not subject to epistemic luck. Now according to anti-luck epistemologists, one fails to know when one's belief is Gettierized, because one's belief is subject to epistemic luck. So, given anti-luck epistemology, if epistemic luck is subject-relative, then it seems that either Blue must be said to have knowledge or anti-luck epistemology fails to explain why subjects in Gettier cases fail to have knowledge. In either case anti-luck epistemology is a failure.

One way of responding to Ballantyne's counterexample is to understand the significance condition on luck in terms of ideal conditions, such as conditions of full rationality. In this case, an event *e* will be significant to an agent S just in case, in conditions of full rationality, S would care whether or not *e* occurred. One might then concede that, while Blue actually doesn't care whether or not he believes the truth that there is a sheep in the field, if he were in conditions of full rationality he would care to believe the truth of the matter. So Blue still meets the significance condition on luck, and it is correct to say that it is lucky for Blue that he believes truly that there is a sheep in the field.

This move fails to fully solve the problem for subject-relative accounts of epistemic luck. It seems that, even in conditions of full rationality, one can be indifferent to whether one believes

extremely trivial truths.[4] For instance, a stranger on the bus tells you that it's his great-great-great grandfather's birthday. In conditions of full rationality, must one care whether one believes this truth? If the correct answer is yes, this isn't obvious.

More importantly, we can still construct scenarios in which it should be insignificant to us that we believe some truth, even if we accept that in conditions of full rationality we will care, all things being equal, to believe the truth that *p*, regardless of what the content of *p* is. Consider the following case. Suppose that if we believe the truth that there is a sheep in the field before us, we will be kept from believing an equally important truth by some evil demon. It follows that our overall epistemic standing will be the same regardless of whether we believe that there is a sheep in the field in front of us. In this case, it seems that even in conditions of full rationality we should be indifferent to whether or not we believe truly that there is a sheep in the field in front of us.

This allows us to construct a variation on Ballantyne's Blue case that spells trouble for understanding epistemic luck in terms of subject-relative luck. Consider the following case.

> Revised Blue: All of the facts of the story hold as above, except for the following. First, it is an objective good for Blue to believe the truth, regardless of what this truth may be. Second, it is not only true that there is a sheep in the field, but it is also true that there is a goat in the field. And third, there is an evil demon who will see to it that Blue will fail to believe the truth that there is a goat in the field if he believes truly that there is a sheep in the field, and vice versa.

In this case, all things considered, it is neither good nor bad for Blue to believe truly that there is a sheep in the field. It follows that it is neither lucky nor unlucky for Blue that he believes truly

[4] See Ballantyne (2011) for a related discussion.

that there is a sheep in the field. Still, Blue's belief that there is a sheep in the field is subject to epistemic luck. His belief is Gettierized.

Pace Ballantyne, the upshot of these reflections should not be a rejection of anti-luck epistemology. Rather, it should be that we do not try to understand epistemic luck as a species of subject-relative luck. We should understand epistemic luck as an instance of subject-involving luck. Note that, even in Revised Blue, it seems right to say that *it is a matter of luck that Blue believes truly that there is a sheep in the field*. This is because there is no significance condition on subject-involving luck. Consider again the Absurdist Raffle case. While winning the absurdist raffle lacks (and should lack) any significance for you, it is still a matter of luck that you win the absurdist raffle. Likewise, even if we suppose that in the revised Blue case, whether or not Blue believes truly that there is a sheep in front of him is (and should be) insignificant to him, it still is a matter of luck that he believes the truth.

It follows that we should understand epistemic luck as a species of subject-involving luck. What about moral luck? Consider the following case of so-called resultant moral luck.

> Drunk Driver: Sam decides to have a liquid lunch, and has way too much to drink. Seeing double, somehow he manages to start his car; he drives home on his usual route, past the local elementary school. He does not hit anyone.

This example serves as an example of resultant moral luck, because, while it is clearly beyond Sam's control whether he hits a child on his way home, his avoiding such an accident determines his moral standing. Had he hit a child, we would judge him not only for being a glutton and reckless but also for being a killer.[5]

[5] See Nagel (1976) for a characterization of moral luck along these lines.

I think considerations similar to those given above show that we should understand moral luck as being a species of subject-involving luck. First, it seems that we have said enough to imply that Sam is not in control of avoiding an accident when we say that it is a matter of luck that he did not hit anyone; second, we can construct scenarios where it is intuitively not lucky for Sam that he fails to hit anyone as he drives home but in which he is still subject to moral luck.

On the one hand, consider examples in which one is indifferent to Sam's moral standing.

> Blue Drunk Driver: Sam has a liquid lunch. He is quite drunk. He makes it to his car, somehow, and drives the normal way from the bar to his home, past the elementary school as the children are being dismissed. Somehow he makes it home without hitting any children. But he is so depressed he is indifferent to having hit a child or not. To him, the world is equally bad, whether or not he is branded a killer.

In this case it, seems that Sam is neither lucky nor unlucky to have avoided hitting any children with his car. It simply doesn't matter to him.

To avoid such cases, again, one might appeal to ideal conditions, for example, conditions of full rationality. One might claim that in conditions of full rationality one must care about one's moral standing because this is somehow tied into being a fully rational agent or perhaps because having a good moral standing is an objective good for one's self. Then one might concede that, while in Sam's actual state it is insignificant to him whether he hit a child, in conditions of full rationality, not hitting a child would be significant to him, and so the significance condition for subject-relative luck is met in Blue Drunk Driver.

But, again, appeals to ideal conditions will not solve the problem. Suppose that we reject *extreme moral rationalism*. Someone is an extreme moral rationalist just in case she believes that nothing can be more significant to someone (in conditions of full rationality) than the quality of her moral standing. Extreme moral rationalism is highly controversial. Once we reject extreme

moral rationalism, cases of moral luck that are neither lucky nor unlucky for a subject abound. To construct such cases, make it a matter of luck that an agent S φs where φing bears on his moral standing. Then stipulate (1) that S's φing ensures that S will ψ, and that S's ψing offsets whatever good is gained or evil suffered by S through his φing.

For instance, suppose that there is an absolute prohibition on lying, and that it is an objective good for everyone to tell the truth. In this case, Sam (in conditions of full rationality) prefers to tell the truth, all things being equal. But suppose that it is also just as much an objective good for Sam to not hurt his friend's feelings. Then it is just as important to Sam, in conditions of full rationality, that Sam not hurt his friend's feelings as it is for Sam to tell the truth. Suppose that Sam has been drinking to the point to which he is not in control over what comes out of his mouth; his friend asks Sam if Sam thinks the friend is a talented poet; Sam speaks truly: "Not really, though I think you are a great person," thereby hurting the friend's feelings. In this case, Sam is not in control in regards to his telling the truth, though *ex hypothesi* this bears on his moral standing. So Sam is subject to moral luck. It seems, however, that Sam is neither lucky nor unlucky to have told the truth. After all, it is just as important to him not to hurt his friend's feelings as to tell the truth. In this case, moral luck is best seen as an instance of subject-involving luck rather than subject-relative luck.

Furthermore, even if one accepts extreme moral rationalism, it is still possible to construct these sorts of scenarios. All we need are events of equal moral significance to cancel each other out, and we can construct scenarios where by φing one is subject to moral luck, though one is neither lucky nor unlucky. Consider the following revision of Drunk Driver:

> Drunk Driver*: Sam decides to have a liquid lunch, and has way too much to drink. Seeing double, somehow he manages to start his car; he drives home on his usual route, past the local elementary school. He does not hit anyone. However,

> this ensures that he will do some other action ψ, which is just as bad as hitting a child with a car while drunk. (For instance, not hitting a child inevitably results in Sam taking target practice with his bow and arrow in a crowded area and shooting a child.)

Here it is neither lucky nor lucky for Sam that he avoids hitting a child with his car on the way home from the bar. In this particular scenario, not hitting a child with his car simply ensures that he will perform some other irresponsible activity that results in the injury of a child. So while it is a matter of luck that he didn't hit a child with his car, and this does affect his moral standing, it seems wrong to say that it is lucky or unlucky for him that he didn't hit a child. For while his moral standing is different from what it would have been if he had hit a child, it isn't better or worse.

It follows that we should not understand moral luck as being an instance of subject-relative luck. Rather, we should understand moral luck as an instance of subject-involving luck. In the cases we canvassed above, it is a matter of luck that Sam does some action such that he is not in control of his performing the relevant action, even though this action bears on his moral standing. But the additional question of whether or not he is lucky to perform the relevant action seems beside the point. Moral luck is not an instance of subject-relative luck.

The upshot of this argument is that, in general, our philosophical efforts should focus on giving an account of subject-involving luck. The topic of luck is of interest primarily for the light it promises to shine on other philosophical topics. But we have strong reasons for believing that focusing on subject-relative luck will fail to deliver on this promise, both in the case of epistemology and in the case of ethics. Thus, we should not worry about subject-relative luck. But we have no reason for rejecting the original intuition that understanding luck can provide

insight into epistemological and ethical topics. Instead of focusing on giving an account of subject-relative luck, we should focus on giving an account of subject-involving luck.

2.4 SUBJECT-INVOLVING LUCK AND LACK OF CONTROL ACCOUNTS OF LUCK

In the last section, I argued that philosophers should shift their focus from subject-relative luck to subject-involving luck. In this section I want to argue that once we focus on giving an account of subject-involving luck, then lack of control accounts of luck, or LCALs, become much more plausible.

According to LCALs of subject-relative luck, an event *e* is lucky for S just in case (i) *e* is (sufficiently) outside S's control, and (ii) *e* meets the relevant significance condition. It is easy to transpose LCALs for subject-relative luck so that they give an account of subject-involving luck. According to LCALs of subject-involving luck, it is a matter of luck that S φs just in case S's φing is not sufficiently under S's control. Different LCALs will give different accounts of what it is for an agent's φing to be under her control. I believe that the most promising line is that taken by Wayne Riggs (2009). Riggs suggests that we understand control in terms of the exercise of an agent's causal powers: "One has control over some happening to the extent that the happening is properly considered something the agent has *done*. . . . This imposes two separable requirements. First, the event has to be the product of the agent's powers, abilities, or skills. Second, the event has to be, at least in some attenuated sense, something the agent *meant to do*" (Riggs 2009, 214).

LCALs have been attacked in both directions. On the one hand, Jennifer Lackey has presented the following case to show that LCAL's do not provide a necessary condition for subject-involving luck:

> Ramona is a demolition worker, about to press a button that will blow up an old abandoned warehouse, thereby completing a project that she and her co-workers have been working on for several weeks. Unbeknownst to her, however, a mouse had chewed through the relevant wires in the construction office an hour earlier, severing the connection between the button and the explosives. But as Ramona is about to press the button, her co-worker hangs his jacket on a nail in the precise location of the severed wires, which radically deviates from his usual routine of hanging his clothes in the office closet. As it happens, the hanger on which the jacket is hanging is made of metal, and it enables the electrical current to pass through the damaged wires just as Ramona presses the button and demolishes the warehouse (Lackey 2008, 258).

According to Lackey, it is lucky for Ramona that she blows up the abandoned warehouse: however, blowing up the warehouse seems to be more or less within her control. So LCALs fail to give necessary conditions for subject-relative luck.

Obviously, if Lackey is right, then this example would also show that LCAL's fail to give necessary conditions for subject-involving luck as well. Lackey's example is convincing, however, only if one vacillates between different readings of *Ramona blows up the building*. If one distinguishes between Ramona's blowing up the building and Ramona's blowing up the building at time *t* (the exact time she blew up the building), then it seems right to say that it was a matter of luck that Ramona blew up the building at time *t*. It seems, however, that Ramona's so doing was *not* under her control. On the other hand, that Ramona blew up the building, *tout court*, seems to be under the control of Ramona, since she is a demolition expert. But understood in this way, it is not a matter of luck that Ramona blows up the building. Even if she hadn't blown up the building at time *t*, she would presumably figure out what was wrong and get the job done.

The problem for LCALs for subject-relative luck is not that they fail to provide *necessary conditions* for an event's being lucky for an agent S. Rather, the real problem is that they fail to provide *sufficient conditions* for an event's being lucky for an agent S. There are many events that intuitively are not lucky for us but are both significant for us and beyond our control. For example, intuitively it is not lucky for us that the sun will rise tomorrow. This event is obviously beyond our control, however, and presumably quite significant for us.[6] Or consider the following example from Lackey (2008, 258). It is quite significant for a child that she get fed her dinner; furthermore, it is largely beyond this child's control that she get fed her dinner. Nevertheless, given that her parents reliably feed her dinner, it is wrong to say that it is a matter of luck that she gets fed her dinner tonight. The upshot of these examples is that while simple LCALs might provide necessary conditions for an event's being lucky for an agent, they do not provide sufficient conditions.

Riggs has tried to supplement the simple LCAL account of subject-relative luck so that it might provide necessary and sufficient conditions for an events being lucky for an agent. His account goes as follows: "E is lucky for S iff (a) E is (too far) out of S's control, and (b) S did not successfully exploit E for some purpose, and (c) E is significant to S (or would be significant, were S to be availed of the relevant facts)" (Riggs 2009, 220). Riggs does not explicitly define what it is for an agent S to exploit an event *e*, rather he provides an illustration of it. Consider, for instance, Riggs's case of Indiana Jones and New Jersey Smith.

> New Jersey Smith plans an expedition into the wilds of Africa where certain tribes of Africans with exotic customs were known to live. . . . He proposes the trip to his fellow adventurer Jones, including the specific times that he means to travel. Jones agrees to tag along. As it happens, the particular tribe that lives in the area that Smith and Jones visit has a custom of sacrificing people from outside the tribe on the equinoxes of the year. The autumnal equinox happens to fall during the time that Smith and Jones are in the area, so they are captured and held until that day so they can be sacrificed. When the day of the autumnal equinox dawns, the tribe readies the captives for

[6] For this example see Latus (2003, 467) and Pritchard (2005, 127).

> sacrifice at midday. As the tribesmen approach to kill them . . . there is a total eclipse of the sun. The members of this tribe always take such exotic natural occurrences to signal the anger of their gods at them for whatever they happen to be doing at the moment. Consequently, they set the captives free (Riggs 2009, 216).

Riggs has us suppose that Indiana Jones knows all along that the tribesmen would likely capture Jones and Smith, and that in response to the eclipse the tribesmen would set their captives free. Jones's knowledge of this figured into his practical plans all along. Smith, unlike Jones, is ignorant of these facts, and so cannot incorporate them into his practical reasoning.

According to Riggs, the occurrence of the eclipse is lucky for Smith but not lucky for Jones. This is because Jones, unlike Smith, exploits the occurrence of the eclipse.

> What seems to distinguish Jones from Smith, and makes Smith lucky to be alive but not Jones, is not that Jones *knew* about the eclipse and whatnot, but that he *exploited* those facts to his own advantage. That is to say, he took them into account and planned a course of action that assumed that those things would occur. And the outcome that resulted, his survival, was a consequence of his having taken account of and exploited those facts. . . . The eclipse was a matter of luck for Smith because it was both out of his control and unexploited by him. The eclipse was not a matter of luck for Jones because, though it was out of his control, he nonetheless exploited its occurrence to procure his survival (Riggs 2009, 218).

The relation of exploitation, then, seems to be something like the following:

> Exploitation: S exploits an event *e*, just in case, (i) before *e* occurs, S uses the belief that *e* will occur in some successful bit of practical reasoning, and (ii) S is justified in so doing.

Riggs's LCAL of subject-relative luck does provide the right results in a number of cases. On Riggs's account, it is not a matter of luck for us that the sun rose this morning, since we all presumably exploited this fact in making our plans for the day. Yesterday, as we engaged in our practical reasoning we took it for granted that the sun would rise today; furthermore, we were justified in so doing. So according to Riggs's account, it is not lucky for us that the sun rose today.

While Riggs's account is an improvement upon previous LCALs of subject-relative luck, it still gets a number of cases wrong. Suppose I've been in a coma for the past ten years.

Presumably, I did not exploit the fact that the sun would rise today. Still, it would not be a matter of luck for me that the sun rose this morning. Or again, take Lackey's case of a young child who is fed dinner by her parents. We might suppose that the child's practical reasoning does not extend more than several minutes into the future, and is always directed to problems that are at hand. We can suppose, then, that the child never assumes that her parents will feed her dinner in any of her practical reasoning. The child, then, fails to exploit the fact that her parents will feed her dinner. Still, it is not lucky for her that she is fed her dinner. So Riggs's account fails to provide sufficient conditions for an event's being lucky for us.[7]

Thankfully, once one turns from giving an account of subject-relative luck to giving an account of subject-involving luck, these sorts of counterexamples are beside the point. When we give an account of subject-involving luck, we are not interested in accounting for why some event is lucky for some agent. Rather, we are interested in accounting for why it is a matter of luck that some agent φs. As a result, once we focus on subject-involving luck, issues such as whether or not it is lucky for an individual that the sun rose today fall out of the picture. As a result, LCALs of subject-involving luck are more attractive than LCALs of subject-relative luck.

[7] Given that the counterexamples to LCALs depend on events that are (1) out of our control and (2) counterfactually robust, such that they occur in all or most nearby possible worlds, couldn't LCALs be saved by adding counterfactual conditions like those that Pritchard (2005, 2007) has proposed? Adding such conditions likely will save LCALs from the claim that they fail to provide sufficient conditions for an event's being lucky relative to an agent S. However, they open LCALs to the objection that they fail to provide necessary conditions for an event's being lucky relative to an agent S. So Lackey (2008) and Hiller and Neta (2007) have provided examples of events which are lucky for an agent but which are counterfactually robust. See Riggs (2009, 207–14) for further discussion.

2.5 OBJECTIONS

Before concluding, I would like to consider two objections.

Objection 1

You have argued that we should not understand instances of epistemic luck and moral luck primarily as instances of subject-relative luck, because there are cases of epistemic luck and moral luck that fail to meet the significance condition. As a result you argue that philosophers should focus their energy on giving an account of subject-involving luck. But this inference might be resisted. It seems that there are instances of epistemic luck and moral luck that are not instances of subject-involving luck. Consider the following case, which is modified from Turri (2011):

> Holmes wants more than anything in the world to boost Watson's confidence as a detective. To do this, he wants to make sure that Watson forms as many true beliefs as possible, without directly intervening in Watson's thought process. Watson is currently investigating a crime scene. Given the evidence that Watson is slowly gathering, he will be justified in believing that the perpetrator of the crime has a limp. Holmes sees that Watson missed an important piece of evidence that shows conclusively that Mr. Plumb, who walks just fine, is the perpetrator. Holmes does not want Watson to get discouraged, so he runs out, finds Mr. Plumb, and kicks him in the knee, ensuring that Mr. Plumb will have a limp. Meanwhile, Watson finishes gathering the evidence and forms the true belief that the perpetrator has a limp (Turri 2011, 5).

Intuitively this is a case of epistemic luck. But it seems that it isn't a case of subject-involving luck. After all, given Holmes's desires and abilities, it isn't a matter of luck that Watson believes truly that the perpetrator has a limp.

Or again, consider a case in which Sam drives home hopelessly intoxicated through busy streets. Unbeknownst to him, his car can be operated by remote control by a friend of his who can observe exactly what the car is doing; the friend is quite good at controlling the car remotely;

the friend will take control if Sam shows signs of being about to hit someone. Here, it is not a matter of luck that Sam avoids hitting someone. After all, his friend will see to it that this occurs. Still, Sam's not hitting anyone is a matter of moral luck.

The upshot of these examples is that you have not provided enough reason for philosophers to shift their attention from providing accounts of subject-relative luck to providing accounts of subject-involving luck. Rather, philosophers concerned with epistemic luck and moral luck should think of these phenomena as sui generis.

Reply to Objection 1

I want to argue that these instances show that there is a certain ambiguity in the phrase "it is a matter of luck that S φs." In one sense, I am willing to grant that it is not a matter of luck that Watson believes truly that the perpetrator has a limp, or that Sam avoids hitting anyone on his way home. After all, there is someone ensuring that this is the case. But in another sense, I want to say it obviously *is a matter of luck* that Watson believes the truth and that Sam arrives home safely. To accommodate these conflicting intuitions, we simply need to recognize that there is a kind of relativity in ascriptions of subject-involving luck. Relative to Watson's and Sam's powers, abilities, and skills, it is a matter of luck that Watson believes truly and Sam arrives home safely. But relative to the powers, abilities, and skills of Holmes and Sam's friend, it is not.

Objection 2

You argue that LCALs become much more attractive once we focus on subject-involving luck. Your basic argument is that counterexamples that plague LCALs of subject-relative luck don't bear on LCALs of subject-involving luck. But this doesn't seem to be true. Consider again Lackey's case of the young girl who is regularly fed dinner. Consider the following event. The

young girl is fed dinner tonight. Now, given that the girl's parents regularly feed her, this event is not a matter of luck. However, it is outside the girl's control that she is fed tonight. So according to LCALs of subject-involving luck it is a matter of luck that the girl is fed tonight. Nothing has been improved.

Reply to Objection 2

LCALs of subject-involving luck attempt to explain what it is for it to be a matter of luck that S φs. "S's φing" here is properly understood as some act of S's, that is, something that S *does*. But in the case mentioned above, "getting fed" is not something the girl is properly said to do. Rather it is something that happens to her. LCALs of subject-involving luck, then, are interested in explaining what it is for it to be a matter of luck that the girl is fed, only in so much as her being fed is an act of her parents or someone else feeding her. And given that it is under her parents' control that the girl is fed, LCALs will state that it is not a matter of luck that her parents feed her.

2.6 CONCLUSION

In this essay I have argued that both epistemic luck and moral luck are best understood as instances of subject-involving luck. As a result, philosophers who believe that understanding luck in general can illuminate issues in epistemology and ethics should focus on understanding subject-involving luck. I have also argued that LCALs become more attractive once we shift our focus. Of course, more work needs to be done. On the one hand, more needs to be said concerning the relativity of ascriptions of subject-involving luck. On the other hand, more needs

to be said concerning what it is for S to have control over his φing. I have only briefly touched upon these matters. I am happy, however, if I have given philosophers interested in luck reason to investigate these issues.

2.7 WORKS CITED

Ballantyne, N. 2011. "Anti-luck Epistemology, Pragmatic Encroachment, and True Belief." *Canadian Journal of Philosophy* 41:485–504.

———. 2012. "Luck and Interest." *Synthese* 185:319–34.

———. 2014. "Does Luck Have a Place in Epistemology?" *Synthese* 1391-1407.

Coffman, E. J. 2007. "Thinking About Luck." *Synthese* 158*:*385–98.

Hiller, A., and R. Neta. 2007. "Safety and Epistemic Luck." *Synthese* 158*:*303–13.

Lackey, J. 2008. "What Luck is Not." *Australasian Journal of Philosophy* 86, no. 2:255-67.

Latus, A. 2003. "Constitutive Luck."*Metaphilosophy* 34:460–75.

Levy, N. 2011. *Hard Luck.* Oxford: Oxford University Press.

Nagel, T. 1976. "Moral Luck." *Proceedings of the Aristotelian Society* 76:136–50.

Pritchard, D. 2007. "Anti-luck Epistemology." *Synthese* 158:277–97.

———. 2005. *Epistemic Luck*. Oxford: Oxford University Press.

Rescher, N. 1995. *Luck: The Brilliant Randomness of Everyday Life.* New York: Farrar, Straus and Giroux.

Riggs, W. "Luck, Knowledge, and Control." In *Epistemic Value,* edited by A. Haddock, A. Millar, and D. Pritchard, 204–21. Oxford: Oxford University Press.

Statman, D. 1991. "Moral and Epistemic Luck." *Ratio,* new series, 4:146–56.

Steglich-Peterson, A. 2010. “Luck as an Epistemic Notion.” *Synthese* 176:361–77.

Turri, J. 2011 “Manifest Failure: The Gettier Problem Solved.” *Philosophers Imprint* 11:1-11

3.0 SAFETY, GETTIER CASES, AND ANTI-LUCK EPISTEMOLOGY

3.1 INTRODUCTION

It is common to hold that luck, or at least a certain kind of luck, is incompatible with knowledge. In particular, it is thought that if it is a matter of luck that one's belief is true (given how one has formed her belief) then one cannot have knowledge (Cf. Pritchard 2005). We can call this knowledge precluding luck *epistemic luck*.[8] This is the kind of luck that is supposed to be at play in Gettier cases.

In the recent literature, accounts of epistemic luck have fallen into two main kinds. On the one hand, there are *virtue-theoretic* accounts of epistemic luck. These accounts state that S's belief that p is subject to epistemic luck just in case S does not believe truly that p because her belief that p is produced by the relevant epistemic virtue (cf. Sosa (2007), Greco (2012), Riggs (2014)). On the other hand, there are *safety accounts* of epistemic luck. A belief is subject to epistemic luck, according to the safety account, just in case one's belief is true but not safe.

Various formulations of what it is for a belief to be safe have been spelled out in the literature. Sosa defines safety as follows:

> Call a belief by S that p "safe" iff: S would believe that p only if it were so that p. (Alternatively, a belief by S that p is "safe" iff: S would not believe that p without it being the case that p: or, better, iff: as a matter of fact, though perhaps not as a matter of strict necessity, not easily would S believe that p without it being the case that p (Sosa 1999, 142).

Pritchard on the other hand, gives the following definition of safety (2007, 2009):

[8] Some authors (cf. Pritchard 2005) prefer to speak of knowledge precluding luck as veritic luck, and to treat veritic luck as a species of epistemic luck. Since the distinctions that these authors wish to highlight are not important for this paper, I will simply use the term epistemic luck.

> *S's* belief is safe *iff* in nearly all (if not all) near-by possible worlds in which *S* continues to form her belief about the target proposition in the same way as in the actual world the belief continues to be true (Pritchard 2007, 283).

While these formulations are different in important ways, it is important to remember that they seek to make explicit the same notion: the notion of a belief's being securely or safely true, i.e., the notion of a belief's being free from danger of cognitive failure.

Pritchard has been by far the most vigorous defender of a safety account of epistemic luck (see his 2005, 2007, 2009, 2012, 2014, and 2015), but the safety account has been defended either explicitly or implicitly by a number of other philosophers. (See for instance Hawthorne (2004) p. 57, n 17), Carter (2010), and Church (2010).)

The safety account of epistemic luck has a number of attractive features. First, it seems that we can give a general account of luck in terms similar to safety. On this modal conception of luck, it is a matter of luck that an event e takes place, just in case, given initial conditions C, it could have easily happened that the event does not take place (cf. Pritchard 2005, 2014)[9]. So the safety account of epistemic luck can be seen as a special case of a modal account of luck. Second, the safety account of epistemic luck is not too demanding. We can compare the safety account of epistemic luck with a sensitivity account of epistemic luck. According to the sensitivity account, S's belief that p is subject to epistemic luck just in case S would have believed that p, even if p were false. The sensitivity account of epistemic luck seems to imply that we cannot have inductive knowledge, since in most cases in which we inductively know that p is true we would have believed that p even if p were false. The safety account of epistemic luck, on the other hand, easily accommodates inductive knowledge (cf. Pritchard 2015, 100). Finally, as Goldberg (Forthcoming) notes, the safety account of epistemic luck is a reductive

[9] Pritchard's currently describes an event's being lucky as follows: '[R]oughly, what makes an eent lucky is that while it obtains in the actual worlds there are – keeping the initial conditions for that event fixed – close possible worlds in which this event does not obtain (Pritchard 2014, 599).'

account: it gives an account of epistemic luck without making reference to any epistemic notions.

Despite the safety account's advantages, it has come under attack. The assault on the safety account has been made by means of putative counterexamples. Take for instance, the following case from Hiller and Neta:

> Orenda believes on good evidence that Jones owns a Ford, and forms the belief that Jones owns a Ford or Brown is in Barcelona. As in the original Gettier case, the belief is justified and true but Orenda does not have knowledge: Jones does not own a Ford but Brown is indeed in Barcelona, unbeknownst to Orenda. But let's also assume that it is no mere accident that Brown is in Barcelona: Brown is so constituted (psychologically, financially, and otherwise) that he would only leave Barcelona in the strangest of circumstances.... In nearly all if not all nearby possible worlds, Orenda's belief is true, but she does not have knowledge (Hiller and Neta 2007, 308).

Orenda does not have knowledge because her belief is subject to epistemic luck, give how she formed her belief it seems that it is a matter of luck that her belief is true -- but her belief *seems* to be safe, at least on Pritchard's official definition of safety. For it seems that in nearly all nearby possible worlds in which Orenda forms her belief in the same way, her belief turns out to be true, and so Orenda is at no risk of forming a false belief.

Hiller and Neta's Orenda case, of course, is an analogue of one of Gettier's original examples, *the Jones owns a Ford or Brown is in Barcelona case*. Analogues have been made for other Gettier cases.[10] Consider for instance, Coffman's (2010) Presumptuous Secretary Case.

> An eminent historian, Hank, recently discovered that Abraham Lincoln was born not in 1809 (as most of us think) but in 1806. Hank writes a letter to his friend, Sandy, in which he asserts that Lincoln was born in 1806. When preparing the letter to be sent, Hank's presumptuous new secretary assumes he made a careless mistake about Lincoln's date of birth, and changes the text so that it says Lincoln was born in 1809. Unbeknownst to all, the secretary's printer has just developed the following glitch: when directed to print a '9', some other number besides '9' is randomly selected and printed instead. As luck would have it, Hank's letter gets printed as stating

[10] I set aside Lackey's (2006) Southernmost Barn Case. This is supposed to be a variant of the Red Barn Façade County case. As in the original case, one is unknowingly driving through Red Barn Façade county (a county which contains only one real red barn, and many red barn facades that would trick anyone who wasn't already in the know), and one drives past the only real red barn in the county; seeing the red barn, one forms the belief that there is a red barn in the relevant area. The case is modified however, so that it is extremely unlikely that one would ever drive past any of the red barn facades. I think that safety accounts of veritic luck can easily deal with this case; as there seems to be little reason for denying knowledge. After all, one's belief is justified, true, produced by a reliable belief forming faculty, and such that one could not easily be wrong about it. I have more to say about this case below.

> that Lincoln was born in 1806. When Sandy receives the letter, she consults it for the answer to the question of when Lincoln was born. Sandy comes to believe, justifiedly and truly, that Lincoln was born in 1806…Sandy could not easily have formed (in the same way) a false belief in a [relevantly similar] proposition. It's not as though (e.g.) the text of the letter is now changing every several seconds, or that Sandy's eyesight has deteriorated to the point that she could easily have misread the relevant sentence. We can safely assume that the letter's text has been perfectly stable for a considerable stretch of time; that Sandy's eyesight is perfect; and so on (Coffman 2010, 246).

Likewise, Sandy Goldberg (Forthcoming) presents a modified version of Chisholm's sheep case:

> **Trick-Of-Light** Through an exceedingly rare trick-of-light, the light rays reflecting off one particular bush will generate the appearance of a rock formation to a person located in one particular spot 20 feet away from the bush. The trick of light is a rarity among rarities: it happens only once a millennium, and even then only has its effect when there is a person who happens to occupy that particular spot 20 feet away, and who happens to be looking in that right direction at the time. Now it happens that there *is* a rock formation on the hill. However, given its location behind the bush, this rock formation is not visible to anyone standing at the spot in question. Imagine finally that is just so happens that there is a subject, S, standing at just that point at just the right time looking in just the right direction, and so who has a perceptual experience as of a rock formation located on the hill. (S has no other access to the rock formation; she does not go closer, or move to get a different orientation.) What is more, the rock formation in question has always been at that spot: it is part of a largely buried rock mass that has been at this site for millions of years, and its presence there reflects fundamental laws of geology. What is more, owing to the laws of nature, the content of the trick-of-light illusion is invariant: whenever (on these very rare occasions) such an illusion obtains, these laws ensure that it comes in the form of an experience *as of a rock formation on the hill.* (No other illusory experience would be compatible with the actual laws of nature.) Now it so happens that S, standing at just the right spot, at just the time when the trick-of-light obtains, and who, looking at the bush in just the right way at just the right time, has a perceptual experience as of a rock formation on that hill. On this basis S comes to believe that there is a rock formation on that hill (Goldberg 2014, 5-6).[11]

In this paper I want to argue for the following claims: first, that these proposed counterexamples to the safety account of epistemic luck do not show that the safety account fails; second, that nevertheless, we have principled reason for believing the safety account of epistemic luck is incorrect.

My paper will proceed as follows. In the first section, I argue that the safety account of epistemic luck can stand up to the putative counterexamples that exist in the literature, and that

[11] Stone (2013) offers a similar case, in which one has a Gettierized belief that meets the safety, sensitivity, and adherence conditions.

we have reason for believing, as things stand, the safety account can deal with any counterexamples that might be provided. However, I think that there is something wrong with the safety-account of epistemic luck. In the second section, I will present a novel argument against safety accounts of epistemic luck that does not rely on any specific counterexample, but on general considerations concerning safety and classic Gettier cases. Finally, I will briefly consider some of the implications that my argument has for work on epistemic luck in general.

3.2 WHY THE COUNTEREXAMPLES DON'T WORK

As noted in the introduction, a number of counterexamples to safety accounts of epistemic luck has been put forward in the literature. The force of such counterexamples can be disputed. On the one hand, in spite of such counterexamples, safety accounts of epistemic luck still appeal to a number of philosophers. On the other hand, even if there are counterexamples to standard safety accounts of luck, perhaps a revised version of the safety account can get around these counterexamples.[12] Indeed, we have strong *prima facie* reason for believing that this is the case.

Consider the following case that Pritchard brings up as a potential problem for safety accounts of knowledge.

> *Mathema*
> Mathema uses a calculator to find out the sum of 12 X 13. As a result, he forms a true belief that 12 X 13 = 156. Unbeknownst to Mathema, however, his calculator is in fact broken and [is] generating 'answers' randomly (Pritchard 2012, 256).

[12] For this same point in a different context, see Bonevac, Dever, and Sosa (2012).

Clearly it is a matter of luck that Mathema's belief is true. But Mathema's belief seems to be safe given Pritchard's definition of safety. After all, there are no possible worlds in in which it isn't true. Pritchard, refuses however, to take the Mathema case seriously as a case in which one's belief is safe. He goes on to say:

> [W]hat we are interested in is… how the agent forms her beliefs in similar circumstances and in response to the same stimuli. These beliefs may be beliefs *that p*, but equally they may be beliefs in distinct propositions… While there is indeed no similar circumstance in which we can imagine Mathema forming a belief that 12 X 13 =156 on the same basis and yet believing falsely, we can certainly imagine lots of similar circumstances in which Mathema forms her belief on the same basis and yet ends up with a false belief, such as the similar situation in which the calculator generates a different result. Mathema's belief is thus unsafe, and hence the safety principle is perfectly able to explain why Mathema lacks knowledge in this case, at least so long as we formulate that principle correctly (Pritchard 2012, 256-257).

The idea here, is that when we consider whether or not S's belief that p is safe, we need to consider not just whether the belief that p when produced by S's method is true throughout nearby possible worlds, but also whether other propositions that could be produced by the method are true in nearby possible worlds. This of course, gets us closer to the core idea of safety, the idea that an individual S's belief is safe just in case S arrived at the belief using a method that didn't put her at serious danger of arriving at a false belief. In the Mathema case it is clear that Mathema did use a method of belief formation that put her in serious danger of arriving at a false belief.

If we keep this in mind and go back to the putative counterexamples given above, it seems that they lose a great deal of their power. Consider for example, Hiller and Neta's Orenda case. There *is* a remarkably similar situation in which she forms her belief in the same way, i.e., performs disjunction introduction on a false but justified belief, and forms a false belief –e.g., she uses disjunction introduction to form the belief that Jones owns a Ford or 2 + 3 =4.

Likewise, consider Coffman's Presumptuous Secretary case. If we describe Sandy as basing her belief regarding Lincoln's birth year on the information in the letter, it seems she

cannot easily form a false belief regarding Lincoln's birth year. However, it seems that there is a remarkably similar situation in which Sandy forms her belief in the same way, i.e., by basing on a letter that was written by Harry's presumptious secretary and printed out on the secretary's faulty printer, in which Sandy's belief is false. So for instance, consider a nearby world in which the secretary's faulty machine printed out an '8' instead of a '6'. In this nearby world, Sandy forms a false belief.

Now Goldberg's Trick-of-Light case is designed to avoid this sort of response. But whether it actually succeeds is doubtful. The trick-of-light case is designed so that one forms one's belief on the basis of an illusion – but the illusion is such that in no nearby world could the illusion lead you to form a false belief. This is because it is supposedly a law of nature that the relevant trick of light will cause you to have an experience of a rock formation being on the hill. Furthermore, it is supposedly a law of nature that the rock formation is located where you have the illusory experience of it being located. So whenever one forms one's belief that there is a rock formation on the relevant basis, i.e., on the basis of one's experience caused by the trick of light, she will always form the true belief that there is a rock formation in front of them.

I think that there is reason to doubt the coherence of Goldberg's case. It seems that there could not be laws of nature that ensure that anyone will have an illusion with a certain type of experiential content. Furthermore, it is doubtful that the laws of nature could guarantee that some rock formation is present at a particular place. This is because laws of nature, on their own, cannot guarantee that any singular fact obtains. But even if we grant that the laws of nature could be as Goldberg describes, the Trick of Light case still plausibly fails as a counterexample to a safety account of epistemic luck. For Goldberg's case to succeed, we need to suppose that the relevant method of belief formation is *not* the following: forming the belief that p on the

basis of an illusory experience. The reasons for this are obvious: once we understand this to be the relevant method of belief formation, there will be all sorts of nearby possible worlds in which one forms false beliefs using the method. But Goldberg needs to explain why we shouldn't understand forming our belief on the basis of illusory experience as being the relevant belief forming process in the Trick-of-Light case.

This problem generalizes. The formula for building counterexamples to safety accounts of epistemic luck goes as follows: first, take a generally reliable method of belief formation, e.g. forming beliefs on the basis of perceptual experience, or forming beliefs on the basis of expert testimony, or forming beliefs on the basis of disjunction introduction performed on justified beliefs etc. etc.; then arrange things so the epistemic agent uses the normally reliable method in circumstances that make the method generally *unreliable*; then place the epistemic agent in circumstances in which they would intuitively not easily form a false belief.

The problem with this way of coming up with counterexamples is that it presupposes that we should understand the method of belief formation in the counterexamples either as the generally reliable method, e.g. forming the belief that p or q on the basis of disjunction introduction and being justified in believing that p, or as a method that is not generally reliable in the sorts of circumstances the epistemic agent is in, but happens to be quite reliable given special characteristics of the actual case, e.g. forming the belief that p or q on the basis of disjunction introduction and a false belief that p, in circumstances in which q is guaranteed to be true. But no principled reason is given for making this assumption. And it is absolutely required for the counterexamples to be successful. For if we consider the relevant method to be, e.g. forming a disjunctive belief on the basis of a justified but false belief and disjunction introduction, it seems

that one's belief is not safe, since in normal conditions using such a method puts one at risk of cognitive failure.

Furthermore, in all of the counterexamples that have been given, it is natural to describe the method used by the epistemic subject in a way that stresses its unreliability/lack of safety. So for instance, it is natural to describe Orenda's method of belief formation as making a deductive inference from a false premise; it is natural to describe Sandy's method of belief formation as believing on the basis of a randomly generated figure; and it is natural to describe one's method of belief formation in the Trick of Light case as forming one's belief on the basis of an illusion.

In effect, all of the counterexamples in the literature take advantage of the generality problem. When an epistemic agent forms a single belief that p, this belief can be understood as being produced by any variety of belief-formation method types. What the putative counterexamples show is that there are cases in which some of these types of belief forming methods seem to be such that they produce safe beliefs and others aren't. For the counterexamples that have appeared in the literature to be true counterexamples, though, it has to be shown that the relevant types of belief formation methods are those that produce safe beliefs.

It is hard to see how this can be done without settling the generality problem. I wait eagerly for this to be done. I think, however, that as things stand, the putative counterexamples merely put pressure on the safety account. To be complete, the safety account needs to explain why the relevant types of belief formation methods in the putative counterexamples are the ones that don't produce safe beliefs.

Of course, one might wonder how much pressure this really puts on the safety account. The generality problem is a problem for any sort of reliabilist epistemology that makes reference to methods/processes/abilities involved in belief formation. And it is the sort of problem that

philosophers attracted to reliabilist epistemology have been more or less happy to live with. So it seems that in the end, the putative counterexamples to the safety account of epistemic luck fail to give us any new reasons for rejecting the safety account. The safety account is a reliabilist account of epistemic luck, at least on a broad construal of reliabilism, and all such accounts all have to deal with the generality problem.

The upshot of these reflections, then, is that we need more than counterexamples to show us that safety accounts of epistemic luck fail. Rather we need a principled reason for thinking that safety accounts of luck fail. This is what I intend to provide in the following section.

3.3 WHY SAFETY ACCOUNTS OF EPISTEMIC LUCK DON'T WORK

The basic argument I make in this section is as follows:

> If the safety account of epistemic luck is correct, then it follows that there cannot be safe, Gettierized beliefs.
>
> But there can be safe Gettierized beliefs.
>
> Therefore, the safety account of epistemic luck is not correct.

The first premise I think should be uncontroversial. It is true that here is some controversy over what exactly counts as a Gettier case. Are Gettier cases simply instances of justified true belief that are not knowledge? Or must they have the double luck structure, described by Linda Zagzebski (1994)? Are Red Barn Façade and other unpossessed defeater cases Gettier cases? In what follows, it is not necessary that we answer all of these questions. We can allow simply that it is a necessary condition for being in a Gettier case, that one's belief

is justified (in the relevant sense), true, and not an instance of knowledge. We can then take Gettier's original cases to be the paradigmatic examples of a Gettier case. Any further case will be a Gettier case, if it sufficiently resembles Gettier's original cases.

It should be clear that Gettier cases are paradigmatic instances of epistemic luck. As several commentators have noted, one of the important lessons that Gettier taught us is that knowledge is incompatible with epistemic luck (cf. Church 2010, 38; Dancy 1985, 134; Pritchard 2005, 4-5; Zagzebski 1999, 99-101). So we have, for instance, Black (2011) saying in a recent text book:

> Gettier's counterexamples to the tripartite account of knowledge, according to which to know is to have a justified true belief, involve epistemic agents whose beliefs are true simply as a matter of luck. These counterexamples are widely… taken to show that an epistemic agent can be justified in holding the true belief that *p* but nonetheless fail to know that *p*. It has thus become a tenet of orthodox epistemology that no belief, not even if it is true and we are justified in holding it, can amount to knowledge if it is true simply as a matter of luck (Black 2011, 187).

Indeed, anti-Gettier epistemology has been spoken of simply as anti-luck epistemology in the literature (cf. Madison 2011). So we are justified in assuming that if S's belief that p is Gettierized then it is subject to epistemic luck, and in thinking that if the safety account of epistemic luck is correct there cannot be safe Gettierized belief.

The second premise, then, is the crucial one. Now, in the past section I argued that none of the proposed counterexamples to the safety account of epistemic luck were effective, and that, until we are given a principled account of method individuation, no counterexamples *can* be effective. Of course, all of these putative counterexamples were supposed to be examples of a safe Gettierized belief. So my argument that there can be safe Gettierized belief will not proceed by way of example. Instead, I need to give principled reason for believing that there can be safe Gettierized beliefs. How can this be done?

One way is by invoking a weakened form of Hume's Dictum (WHD). Consider the relation of partial constitution that different states of affairs or different kinds of events can have towards each other. A kind of event E partially constitutes another kind of event F just in case what it is for an event of type F to take place is (in part) for an event of type E to take place. If event type E and F are such that neither partially constitutes the other, then we can say that they are wholly distinct types of events. According to the weakened form of Hume's Dictum I endorse, the following is true:

> (WHD) If two types of contingent events (or states of affairs) are wholly distinct then we have (ceteris paribus) sufficient reason for denying there is a necessary connection between these types of events.

A few words are in order in regards to (WHD). It should be noted that (WHD) is an epistemic and not a metaphysical principle. (WHD) is compatible with there being a necessary connection between two contingent and wholly distinct types of events; (WHD) merely tells us that as a default we should believe that it is possible for wholly distinct types of events to occur without each other. In this way (WHD) is similar to McPherson's (2012) Modest Humean principle which states that any commitment to brute necessary connections between discontinuous properties counts significantly against a view (McPherson 2012, 217).

(WHD) then, is a relatively weak principle, and in itself should not cause trouble for the safety-theorist of epistemic luck.[13] But we can use it to argue for the possibility of safe-Gettierized beliefs.

To see this consider one of Gettier's original cases:

> Let us suppose that Smith has strong evidence for the following proposition:
>
> (f)... Jones owns a Ford...

[13] Or for that matter, anyone who wishes to put forward the safety condition as an anti-Gettier condition for knowledge.

> Let us imagine, now, that Smith has another friend, Brown, of whose whereabouts he is totally ignorant. Smith… constructs the following [proposition]…
>
> (h) Either Jones owns a Ford, or Brown is in Barcelona.
>
> [This proposition] is entailed by (f). Imagine that Smith realizes the entailment of [this proposition] he has constructed by (f), and proceeds to accept… (h) on the basis of (f)... But now imagine that two further conditions hold. First, Jones does not own a Ford, but is at present driving a rented car. And secondly... the place mentioned in proposition (h) happens really to be the place where Brown is (Gettier 1963, pp. 122-123).

The crucial aspects of the case are as follow.

(i) Smith is justified in believing that Jones owns a Ford.
(ii) It is false that Jones owns a Ford.
(iii) Given that Smith is justified in believing that Jones owns a Ford, then Jones is justified in believing Jones owns a Ford or Brown is in Barcelona on the basis of believing that Jones owns a Ford and disjunction introduction.
(iv) Smith bases his belief that Jones owns a Ford or Brown is in Barcelona only on his believing that Jones owns a Ford and disjunction introduction.
(v) It is true that Jones owns a Ford or Brown is in Barcelona.

It is clear that none of the crucial elements in the case need to be understood in terms of Smith's disjunctive belief being unsafe. Indeed, it would be bizarre to try to give an account of any of these elements in such terms. Consider, for instance, an account of what it is for Smith to have a justified belief that Jones owns a Ford in terms of Smith's having an unsafe belief that Jones owns a Ford or Brown is in Barcelona. Such an account is not even intelligible. Nor could we give an account of what it is for the conjunction of states of affairs described in (i)-(v) to obtain in terms of a lack of safety in Smith's disjunctive belief. What it is for this conjunction of states of affairs to obtain is simply that each of the conjuncts obtains.

Nor should we say that part of what it is for Smith to have an unsafe belief that Jones owns a Ford or Brown is in Barcelona is for (some of) the states of affairs described in (i)-(v) above to obtain. Smith has an unsafe belief just in case Smith forms this belief in a way that puts him at serious risk of forming a false belief. But Smith can do this, of course, without any of the states of affairs described in (i)-(v) above obtaining. Perhaps, for instance Smith believes that

Jones owns a Ford or Brown is in Barcelona merely on a whim. So formed, his belief would be unsafe, but perhaps none of the facts listed in (i)-(v) would obtain.

Thus, the state of affairs that is Smith's having an unsafe disjunctive belief that either Brown is in Barcelona or Jones owns a Ford and the states of affairs described in (i)-(v) are not partially constitutive of each other. In other words they are wholly distinct. Given (WHD), it follows that we have sufficient reason for believing that there are states of affairs in which (i)-(v) obtain and Smith's disjunctive belief is not unsafe.

All of this needn't cause problems for the safety theorist, except that the conjunctive state of affairs (i)-(v) seem to *entail* that Smith's belief is Gettierized. That is to say, it seems that in all possible worlds in which (i)-(v) obtain, regardless of what else obtains in these worlds, Smith's belief that either Jones owns a Ford or Brown is in Barcelona is a Gettierized belief. After all, we have understood Gettierized beliefs to be beliefs that are sufficiently similar to the beliefs described in the original Gettier cases, and it is plausible to suppose that any case in which (i)-(v) obtain is sufficiently similar to the original Gettier cases to count as a Gettier case. Of course, I cannot prove this in a non-question begging way against the safety theorist. But I can offer it as a challenge to the safety theory. If the safety theory of epistemic luck is correct, there should be some further states of affairs that can be added (i)-(v) that would de-Gettierize Smith's belief. It is up to the safety-theorist to provide them. However, it is hard to see what these further states of affairs might be.[14]

As things stand we have reason to believe Smith's having an unsafe belief is not constitutive of the conjunctive state of affairs described by (i)-(v) above, and these facts together

[14] Of course it would be question begging for the safety-theorist to claim that all that needs to be added to (i)-(v) is that Smith's belief is safe.

entail that one's belief is Gettierized. So we have reason for thinking that Smith's disjunctive belief might be safe when facts (i)-(v) above obtain, i.e., when Smith's disjunctive belief is Gettierized.

3.4 ANTI-LUCK EPISTEMOLOGY WHERE NOW

If my argument is correct, then we should believe that there are safe, Gettierized beliefs. As a result, we should reject the safety-account of epistemic luck.

One might reply that at best my argument shows that while certain Gettierized beliefs are paradigmatic instances of epistemic luck, not every Gettierized belief is subject to epistemic luck. That is, one might deny the first premise of my argument. But I don't think this is the way to go. The existence of safe Gettierized belief would give us reason for thinking that not every Gettier case is a case in which epistemic luck is at play, *if* we were forced to accept the modal account of luck as the one, true account of luck (cf. Pritchard (2005), (2014)). But there is reason to believe that there are many different senses in which something can be a matter of luck. (Cf. Riggs (2014), Greco (2012).) What the existence of safe Gettierized beliefs shows us, then, is that the modal account of luck is insufficient for fully capturing the notion of epistemic luck.

A better response for the safety theorist is to claim that we need to incorporate safety conditions into a fuller account of epistemic luck. This is the tack that Setiya (2013) and Schafer (2014) seem to suggest. Setiya, for instance, argues that safety conditions need to be supplemented with at least explanatory conditions to give a full account of epistemic luck. Shafer (2014) on the other hand, suggests that safety conditions will need to be supplemented

with virtue/credit conditions. It should be noted that this view is not far removed from Pritchard's (2012) *anti-luck virtue epistemology*, except that instead of understanding there to be two distinct conditions on knowledge – an anti-luck condition and an ability condition – there are two distinct conditions on non-accidentally true belief – safety conditions and credit conditions.

There is still room to doubt, however, that safety conditions have even a mitigated role to play in anti-Gettier epistemology. For the safety condition to have a role to play in anti-Gettier epistemology, there need to be Gettier cases which can be spelled out in terms of events or states of affairs that are partially constituted by an epistemic agent's belief being unsafe, and cannot be ruled out by whatever conditions rule out *safe* Gettier cases. Otherwise there is no work for the safety condition to do.

It is unclear however, if there are any such cases. The most plausible candidates are so-called 'dangerous Gettier cases', i.e., red-barn façade cases and their analogues. On the one hand, it is generally accepted that virtue and credit conditions struggle to rule out knowledge in red barn façade cases (cf. Sosa 2007, Lackey (2009), Greco (2012) *inter alia*). On the other hand, it can be argued that when we make one's belief safe in red barn façade cases, this in effect de-Gettierizes them. To see this, consider Lackey's (2006) *Southernmost Barn* case:

> [W]hile entering a Midwestern farming community on her cross country drive, Janice looked at the first barn that she saw, which was on the southernmost end of the field, and formed the corresponding belief 'There is a barn'. As it happens, the barn she saw is the only real one, surrounded by barn facades that members of this community have placed in the field in order to make their town appear prosperous. However, as a matter of strict and unwavering policy, the members of this community always place their only real barn on the southernmost end of their land, since this is where traffic first enters their town. Moreover, thirty years earlier, Janice had lived in a house on the southernmost end of this field in the precise location of the one real barn. Because of her deep interest in her childhood roots combined with the brief period during which she can safely take her eyes off her driving, she would invariably have looked at the only particular place in the field where the real barn exists (Lackey 2006, 288).

In this case it seems that Janice does have knowledge. After all, she has a safe belief that is creditable to her epistemic abilities. What then is lacking? The key factor then, in determining whether safety conditions are necessary for anti-Gettier epistemology, rests on whether we have knowledge in dangerous Gettier cases generally. A growing number of epistemologists have argued that we do have knowledge in these cases (cf. Hawthorne and Szabo (2004), Sosa (2007), Turri (2011), for some empirical work testing folk intuitions see Colaco et. al. (2014)). If they are right, then anti-luck epistemology should not be safety-based epistemology.

3.5 WORKS CITED

Bonevac D., Dever, J. and Sosa, D. (2011). "The Counterexample Fallacy," *Mind* 120:1143-1158

Black, T. (2011). "Anti-Luck and Modal Epistemology," in *The Routledge Companion to Epistemology* eds. Bernecker and Pritchard (New York: Routledge, pp. 187-198).

Bogardus, T. (2014). "Knowledge Under Threat" *Philosophical and Phenomenological Research* 88:289-313.

Carter, J. (2010). "Anti-Luck Epistemology and Safety's Recent Discontents," *Philosophia* 38:517-532

Church, I. (2010). "Getting 'Lucky' with Gettier," *The European Journal of Philosophy* 21:37-4

Coffman, E. (2010). "Misleading Dispositions and The Value of Knowledge" *The Journal of Philosophical Research* 35:241-258.

Colaco, D., Buckwalter W., Stich S. and Machery E. (2014). "Epistemic Intuitions in Fake Barn Thought Experiments" *Episteme* 11:199-212.

Dancy, J. (1985). *An Introduction to Contemporary Epistemology*, (Oxford: Basil Blackwell)

Gettier, E. (1963). “Is Knowledge Justified True Belief?” *Analysis* 23:121-123

Goldberg, S. (Forthcoming) “Epistemic Entitlement and Luck” *Philosophy and Phenomenological Research*.

Greco, John. (2012). “A (Different) Virtue Epistemology” *Philosophy and Phenomenological Research* 85: 1-26

Hawthorne, J. and Szabo, T. (2005). “The Real Guide to Fake Barns: A Catalogue of Gifts for Your Epistemic Enemies” *Philosophical Studies* 124:331-352.

Hiller A. and Neta, R. (2007). “Safety and Epistemic Luck,” *Synthese* 158:303-313

Lackey, Jennifer. (2009). “Knowledge and Credit ” *Philosophical Studies* 142:27-42.

----. (2006). “Pritchard’s Epistemic Luck” *Philosophical Quarterly* 56:284-289.

Madison, B.J.C. (2011). “Combatting Anti Anti-Luck Epistemology” *The Australasian Journal of Philosophy*, 89:47-58.

McPherson, T. (2012). “Ethical Non-Naturalism and the Metaphysics of Supervenience” in *Oxford Studies in Metaethics Vol. 7* ed. R. Shafer-Landau. (Oxford: Oxford University Press) 205-234.

Pritchard, D. (2007). “Anti-Luck Epistemology” *Synthese* 158:277-297

----. (2015). “Anti-Luck Epistemology and the Gettier Problem” *Philosophical Studies*

----. (2012). “Anti-Luck Virtue Epistemology”, *The Journal of Philosophy*

----. (2005). *Epistemic Luck* (Oxford: Oxford University Press)

----. (2009). “Safety Based Epistemology: Whither Now?” *The Journal of Philosophical Research* 34:33-45.

----. (2014). “The Modal Account of Luck” *Metaphilosophy* 4-5:594-619.

Riggs, W. (2014). “Luck, Knowledge, and Mere Coincidence” *Metaphilosophy*: 4-5: 627-639.

Setiya, K. (2012). *Knowing Right from Wrong* (Oxford: Oxford University Press).

Schafer, K. (2014). “Knowledge and Two-Forms of Accidental Truth” *Philosophy and Phenomenological Research* 89: 373-393.

Sosa, E. (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Vol. 1*. (Oxford: Oxford University Press)

----. (1999). “How to Defeat Opposition to Moore” *Philosophical Perspectives* 13:141-153.

Stone, J. (2013). “Unlucky Gettier Cases” *The Pacific Philosophical Quarterly* 94:421-430.

Turri, J. (2011). “Manifest Failure: The Gettier Problem Solved.” *Philosophers Imprint* 11:1-11

Zagzebski, L. (1994). “The Inescapability of Gettier Problems” *The Philosophical Quarterly* 44:65-73.

--- (1999). “What is Knowledge?” in Greco and Sosa, *The Blackwell Guide to Epistemology* (Oxford: Blackwell Publishing)

4.0 AGAINST MIXED EPISTEMOLOGY

4.1 INTRODUCTION

Many philosophers believe that there is a *safety condition* on knowledge. The exact formulation of the safety condition is controversial, but the basic idea is that if S knows that p, then S forms her belief that p in such a way that she is not at risk of forming a false belief (for this formulation of the core idea of safety, cf. Bogardus (2014)).

There are also many philosophers who believe there is an *ability condition* on knowledge. Again, the exact formulation of the ability condition is controversial, but the basic idea is that if S knows that p, then S's true belief is creditable to a cognitive ability of S's (cf. Sosa (2007), Schafer (2014)).

Philosophers have tried to give a reductive account of knowledge solely in terms of the safety condition (e.g. Pritchard 2005, 2007) or solely in terms of the ability condition (e.g. Sosa (2007), Greco (2010, 2012), and Turri (2011, 2015). Following Pritchard (2012) we can respectively call these different reductive accounts *pure anti-luck accounts* and *pure virtue accounts*. In turn we can call any reductive account of knowledge that appeals to both safety conditions and ability conditions a *mixed account of knowledge*.

Recently, Duncan Pritchard (2012) has argued for a mixed account of knowledge, defending what he calls *Anti-Luck Virtue Epistemology* (ALVE). According to ALVE:

> *S* knows that *p* if and only if *S*'s safe true belief that *p* is the product of her relevant cognitive abilities (such that her safe cognitive success is to a significant degree creditable to her cognitive agency) (Pritchard 2012, 273).

Christoph Kelp (2013) defends his own mixed account of knowledge, which he calls the *Safe-Apt* view of knowledge. Says Kelp:

> The core thesis of virtue epistemology is that knowledge involves cognitive success that is due to cognitive competence. More specifically… One knows that *p* only if one believes *p* and the acquisition and/or retention of one's true belief that *p* is (sufficiently) due to cognitive competence or ability… I want to suggest that combining… an additional safety condition with a Sosa-style competence-manifestation account of the 'because' relation [in the core thesis of virtue epistemology] can avoid a number of problems that beset its competitors (Kelp 2013, 266).

John Turri (2011) has considered another mixed account of knowledge. According to Turri, knowledge might be *ample* belief; i.e safe belief in which the safety of the belief manifests the believer's cognitive ability.[15]

There are two reasons that philosophers interested in reductive accounts of knowledge have been attracted to mixed accounts of knowledge. First, both pure safety accounts and pure ability accounts struggle with well-known counterexamples. However, the counterexamples that plague pure safety accounts do not seem to affect pure ability accounts, and likewise, the counterexamples that plague pure ability accounts do not seem to affect pure safety accounts. The thought is that by combining both safety and ability conditions we can give an extensionally adequate reductive account of knowledge. Second, philosophers like Pritchard and Kelp argue that given the function of knowledge ascriptions, we should expect that there would be distinct safety and ability conditions on knowledge.

In this paper I want to focus on the first motivation for mixed accounts of knowledge, i.e., that while pure safety and pure ability accounts struggle to achieve extensional adequacy, mixed epistemology does not. I argue that adding an ability condition to the safety condition and *vice*

[15] Turri has since disavowed this possibility (cf. Turri (2015)). It is worth noting that Turri now accepts unequivocally that we have knowledge in Red Barn façade cases.

versa does nothing to help with the putatively problematic cases. This is because both the safety condition and the ability condition are ways of spelling out the intuition that knowledge must come from reliable methods of belief formation in the right way, but the putative counterexamples to both pure safety accounts and pure ability accounts arise because of problems inherent in our notion of reliability.

The paper proceeds as follows. In the first section I will look at the putative counterexamples for pure safety and pure ability accounts that are used to motivate mixed accounts of knowledge. In the second section I will argue that these sorts of cases are equally problematic (or unproblematic) for both pure safety and pure ability accounts. In the third section, I will argue that the putative counterexamples arise because both the safety and ability conditions are attempts to spell out the intuition that knowledge must come from a reliable source, and attributions of reliability are context sensitive.

4.2 THE CASES

Mixed accounts of knowledge are motivated by counterexamples to pure safety accounts of knowledge which are not counterexamples to pure ability accounts and *vice versa*. What are these counterexamples?

First there are so-called dangerous Gettier cases. The most famous of these cases is the Red Barn Façade case. Here is Pritchard's formulation.

> *Barney*
>
> Using his reliable perceptual faculties, Barney non-inferentially forms a true belief that the object in front of him is a barn. Barney is indeed looking at a barn. Unbeknownst to Barney, however, he is in an epistemically unfriendly environment when it comes to making observations of this

> sort, since most objects that look like barns in these parts are in fact barn facades (Pritchard 2012, 251).

Pure virtue epistemologies are supposed to struggle with ruling out dangerous Gettier cases like the *Barney case*. In this case, Barney is said to form the true belief that there is a red barn in front of him using his reliable perceptual faculties; we can further stipulate that lighting conditions etc. are normal. It follows that Barney's true belief that there is a red barn in front of him is attributable to a reliable method of belief formation; furthermore, this reliable method of belief formation *is* a cognitive ability. After all, reliable perceptual faculties are paradigmatic instances of epistemic abilities. Thus, *prima facie*, in the *Barney case* and other similar cases, pure ability accounts must rule that Barney has knowledge. And indeed, many epistemologists that adopt pure ability accounts of knowledge are happy to do so.[16]

Epistemologists defending a mixed account of knowledge assume that, while the ability condition struggles to rule out dangerous Gettier cases, the safety condition easily does so. After all, Barney is at risk for forming a false belief – he could have easily been looking at a barn façade and would have formed a false belief that there was a barn in front of him.

If the safety condition deals easily with dangerous Gettier cases like the *Barney case*, however, it struggles to deal with other sorts of cases in which one's true belief is safe, but safe for the wrong sorts of reasons. Consider for example, Pritchard's *Temp* case.

> *Temp*
>
> Temp forms his beliefs about the temperature in the room by consulting a thermometer. His beliefs, so formed, are highly reliable, in that any belief he forms on this basis will always be correct. Moreover, he has no reason for thinking that there is anything amiss with his

[16] How much of a strike this is against such theories is difficult to assess, because many philosophers, and not just pure virtue epistemologists, now attribute knowledge to subjects in dangerous Gettier cases (cf. Sosa (2007), Turri (2011), Lycan (2006), Millikan (1995) *inter alia* for philosophers who believe that subjects have knowledge in these cases, cf. Szabo and Hawthorne (2005) for some of the difficulties in denying knowledge to those in dangerous Gettier cases).

> thermometer. But the thermometer is in fact broken and is fluctuating randomly within a given range. Unbeknownst to Temp, there is an agent hidden in the room who is in control of the thermostat whose job it is to ensure that every time Temp consults the thermometer the 'reading' on the thermometer corresponds to the temperature in the room (Pritchard 2012, 260).

Prima facie, Temp's belief is safe, but it isn't an instance of knowledge. So pure safety accounts of knowledge are forced to give the wrong verdict regarding the *Temp case*.

Epistemologists giving a mixed account of knowledge assume that the ability condition easily rules out cases like the *Temp case*. The thought is that in the *Temp case*, Temp's belief is not creditable to any method of belief formation that can be called a virtue or an ability. As Pritchard puts it, "[the] correctness [of Temp's beliefs] has nothing to do with Temp's abilities and everything to do with some feature external to his cognitive agency" (Pritchard 2012, 260).

So while the ability condition struggles to rule out cases like the *Barney case*, it does a fine job with the *Temp case*, and while the safety condition struggles with the *Temp case*, it does a fine job with the *Barney case*. Mixed accounts of knowledge take advantage of these putative facts and combine both safety and ability conditions in a reductive account of knowledge.

4.3 WHY THE CASES DON'T MOTIVATE MIXED ACCOUNTS OF KNOWLEDGE

In the previous section I reviewed some cases that are supposed to motivate mixed accounts of knowledge. In this section, I argue that, in fact, these cases do *not* motivate mixed accounts of knowledge. If ability conditions on knowledge struggle to rule out dangerous Gettier cases like the *Barney case*, so do safety conditions; if safety conditions on knowledge struggle to rule out cases like the *Temp case*, so do ability conditions.

First, I will consider dangerous Gettier cases like the *Barney case*. Despite initial appearances to contrary, the safety condition on knowledge struggles to rule out dangerous Gettier cases as instances of knowledge. To see this, consider the following case from Comesaňa (2005) that he brings up as a counterexample to the safety condition on knowledge.

> HALLOWEEN PARTY: There is a Halloween party at Andy's house, and I am invited…. [Andy] hires Judy to stand at a crossroads and direct people towards the house… Unbeknownst to me, Andy doesn't want Michael to go to the party, so he also tells Judy that if she see Michael she should tell him the same thing she tells everybody else (that the party is at the house down the left road), but she should immediately phone Andy so that the party can be moved to Adam's house, which is down the right road. I seriously consider disguising myself as Michael, but at the last moment I don't. When I get to the cross roads, I ask Judy where the party is, and she tells me that it is down the left road (Comesaňa 2005, 397).

In the Halloween case, you have knowledge, but your belief doesn't seem to be safe: i.e., you seem to be at risk of forming a false belief. So it seems that there is no safety condition on knowledge.

The right response to this kind of counterexample, is to *relativize* the safety condition to a method of belief formation. In this case beliefs are not safe *simpliciter*, but are safe *relative* to the method by which they were formed (cf. Pritchard 2007, 2009). Given that you form your belief as to where the party is on the basis of *knowledgeable and sincere testimony*, you are not at risk of forming a false belief. So *pace* Comesaňa your belief in the *Halloween party case* is an instance of knowledge, and it *is* safe. The *Halloween party case* only appears to threaten the safety condition on knowledge because you were at risk of not forming your belief on the basis of *knowledgeable and sincere testimony*. If, however, we don't confuse being at risk of forming a false belief given one's method of belief formation, and being at risk of not using a safe method of belief formation, then counterexamples like *Halloween party case* pose no threat to the safety condition on knowledge (cf. Bogardus 2014, 297-299).

However, when we relativize the safety condition to a method of belief formation, using the safety condition to rule out Red Barn facades and other dangerous Gettier cases becomes problematic. To see this, consider a method m such that using m entails the truth of p. We can call such methods *factive* methods. Properly deducing p from q when one knows that q is an example of such a method. Other factive methods include believing that p on the basis of seeing that p, believing that p on the basis of sincere, competent, and knowledgeable testimony that p, and believing that p on the basis of correctly remembering that p. Every method of belief formation that is a factive method produces *only* maximally safe beliefs. Given that you form your belief on the basis of seeing that p, you are at *absolutely no* risk of forming a false belief that p (or any closely related proposition) since, given that you form your belief on the basis of seeing that p, your belief that p must be true.

We can give a plausible description of the *Barney case* in which Barney forms his belief that there is a red barn in front of him on the basis of *seeing that* there is a red barn in front of him. Under this description, Barney's belief will be safe. For again, given that Barney's belief that there is a red barn in front of him is based on his seeing that there is a red barn in front of him, there is *no* risk that Barney will believe falsely.

If safety is to rule out the *Barney* case then it must be assumed that the method that Barney is using is non-factive. Suppose for the sake of argument that this assumption is justified. Then it seems that the safety condition implies that Barney lacks knowledge. We can, for the sake of illustration, suppose that Barney bases his belief that there is a red barn in front of him on the fact that he has no reason for doubting that he is in normal conditions, and it appears to him that there is a red barn in front of him. (The exact details of the method don't matter for our purposes.) In this case, it seems that Barney's belief that there is a red barn in front of him is

not safe. In many nearby worlds, worlds in which he just happens to be facing one of the many barn facades littering the countryside, Barney forms false beliefs using the same method, and so Barney *is* at risk of forming a false belief given the method of belief formation he is using.

But even on the assumption that Barney's method of belief formation is non-factive, there are problems. Consider the following case.

> **Red Barn County Architect**: Peter is an epistemic thief. At great personal expense, and after years of research and development, he created the Red Barn Façade County Erector Set. Once the erector set was put in place, at the push of button, any rural county could be turned into the epistemic equivalent of Red Barn Façade County. (The erector set was amazing – it could instantly raze all but one of the red barns in the county and set up any number of red barn facades in their place.) One day, it turns out that Barney is driving through a rural county in which Peter had set up his erector set. Of course Barney is unaware of this. He is about to fall into the trap. With Barney approaching a red barn that is to be razed and replaced by a red barn facade, Peter pushes the button! Then the completely unexpected happens. Peter's almost perfectly reliable erector set fails to work. Barney sees the red barn in front of him and forms the belief: there is a red barn in front of me.

Here it seems that Barney obviously has knowledge that there is a red barn in front of him. But in many of the nearest possible worlds in which Barney forms his belief that there is a red barn in front of him on the basis that has no reason for believing that conditions are not normal, and it seems to him that there is a red barn in front of, Barney forms a false belief. So it can seem that in the *Red Barn County Architect case* Barney *is* at risk of forming a false belief if we assume his method of belief formation is not factive.

These reflections show that the safety condition on knowledge cannot deal with dangerous Gettier cases in a simple way. The problem is that we cannot say whether a belief is safe until we have fixed the relevant method of belief formation, but in dangerous Gettier cases (e.g the *Barney case*) and in cases that are almost dangerous Gettier cases (e.g. the *Red Barn Façade Architect case*) we can give alternative descriptions of the relevant methods of belief

formation that make the belief respectively safe or unsafe, where these alternative descriptions are equally plausible.

In effect, we have an analogue to the generality problem for process reliabilism. While various solutions to the generality problem have been put forward, the one that comes most easily to hand is Heller's (1995) *contextualist* or *simple* response. According to Heller there is no fact of the matter whether a given belief is safe or not, as there is no fact of the matter as to what methods of belief formation are relevant for the epistemic assessment of a belief.[17] Rather, there are contextual factors that determine what the relevant methods of belief formation are for deciding whether a belief is safe. In this case, it won't be true that a belief is formed on the basis of some method m absolutely speaking but only relative to some context of knowledge attribution.[18]

Of course, the details of this contextualist treatment of the safety condition are controversial and tough to make out. The important point is that the kind of contextualist maneuver needed by safety theorists to successfully deal with the *Barney case* could also allow credit theorists to give the desired verdict. This is especially clear if we think of the methods of belief formation relevant to credit theorists as being reliable methods of belief formation relative to some environment. Indeed Greco (2010), (2012) appeals to contextualist ideas to allow his pure ability account to deal with dangerous Gettier cases. Once again, the details of this contextualist treatment of epistemic ability will be controversial, but in this the pure ability accounts of knowledge are no worse off than pure safety accounts of knowledge. The upshot of

[17] It should be noted that Heller does not explicitly speak of 'safety' but instead of 'reliability'. *Mutatis mutandis,* his remarks apply equally well to safety.

[18] My appeals to context here are supposed to be neutral between semantic contextualism and subject sensitive invariantism. The point is simply that whether it is true that an individual S forms her belief that p using a method m depends on factors concerning the attributor's circumstances, values, etc.

these reflections, then, is that adding safety conditions to pure accounts of knowledge gives them no real advantage in ruling out cases like the *Barney case*.[19]

Of course, mixed epistemologists might reject Heller's contextualist solution to the generality problem. But in this case, they must provide their own solution to the problem. Until this is done, it is unclear how adding a safety condition to knowledge helps to rule out dangerous Gettier cases.

Next consider cases in which one's belief appears to be safe, but is safe for the wrong reasons, such as the Temp case. Despite initial appearances to the contrary, the ability condition on knowledge fails to rule out the *Temp case* as a case of knowledge in a straightforward manner.

Pritchard believes that Temp fails to meet the ability condition on knowledge, because "[the] correctness [of Temp's beliefs] has nothing to do with Temp's abilities and everything to do with some feature external to his cognitive agency" (Pritchard 2012a, 260). But this is not true. The correctness of Temp's beliefs does seem to be attributable, at least in part, to Temp's cognitive agency. After all, Temp wouldn't form the correct belief if he didn't have the ability to read the thermometer, or if he didn't think that it was rational to trust the thermometer, and instead thought the rational thing to do was to believe that the temperature in the room is anything but what the thermometer read. And it seems that if Temp was so mad as to form any

[19] If my argument is correct, then requiring that one's belief be safe does not help one to rule out dangerous Gettier cases like the *Barney* case. This weighs against both weaker and stronger versions of mixed epistemology. Remember that on the weaker version of mixed epistemology, one's belief must meet the standard virtue/credit conditions for knowledge, and one's belief must meet the safety condition. On the stronger version of mixed epistemology one's belief must meet the safety condition, and its doing so must be attributable to one's epistemic virtue/ability. Of course, the problem is that much more difficult for stronger versions of mixed epistemology. For it is clear that, even if one's belief is safe in the *Red Barn Façade Architect* case, its safety does not seem to be attributable to one's epistemic virtue or cognitive ability.

random belief on the basis of what the thermometer reads, e.g. if he was disposed not only to form beliefs concerning the temperature on the basis of reading the thermometer but was also disposed to form beliefs regarding the price of oil futures on the basis of reading the thermometer, we couldn't so much as consider his belief to be safe. So Temp's forming the correct beliefs regarding the temperature in the room is attributable, in some part at least, to his cognitive agency.

If appeals to the ability condition are to rule out the *Temp case*, they must do so for reasons other than the putative fact that Temp is in no way creditable for his true belief. But it seems that this cannot be done by demanding that *the greater* share of credit for one's cognitive success belong to the knowledgeable subject. As Lackey (2009) and others have noted, this move seems to imply that we cannot have knowledge on the basis of testimony from experts, since in these cases the greater share of credit for one's cognitive success goes to the expert and not to the one who accepts the testimony.

A possibility that has been explored in the literature is that we should understand the ability condition in terms of 'manifestation'. On this view, believing truly that p can be creditable to an agent S, just in case in believing truly that p, S manifests her cognitive ability (Cf. Turri 2011, Sosa 2010). This is exactly the line of thought that Kelp (2013) picks up on. On this way of thinking, Temp's true beliefs regarding the temperature in the room fail to be an instance of knowledge, because Temp's cognitive success in these instances fails to manifest his cognitive ability. We can contrast this with Temp's beliefs about the temperature formed on the basis of reading the thermometer in normal conditions. In this case, Temp's cognitive success *does* manifest his cognitive ability.

Prima facie this manner of dealing with the problem shows some promise, but I think it ultimately fails. Consider that Temp's believing truly in the *Temp case* manifests a disposition of his to get things right regarding room temperature in conditions described by the *Temp case*. If the ability condition is to rule out that Temp has knowledge, then it needs to be able to rule out the possibility that this disposition is a cognitive ability of Temp's.

A possible way of dealing with this problem is to hold that epistemic abilities must be an integrated part of one's cognitive character. Greco (2010) takes this tack in dealing with the so-called problem of strange and fleeting processes, and Pritchard follows suit. Says Pritchard:

> We… need to ask what it is about the particular belief-forming dispositions that qualify as cognitive abilities that makes them knowledge conducive. I take it that as a minimal requirement these belief-forming dispositions should be both reliable and suitably integrated with the agent's other belief-forming dispositions (Pritchard 2012, 262).

But what is it for a belief-forming disposition to be suitably integrated with the agent's other belief-forming dispositions? And why shouldn't we consider the relevant dispositions to get things right in the *Temp case* to be well-integrated belief-forming disposition?

According to Greco cognitive integration is a matter of both a range of outputs and how this range of outputs interacts with other belief-forming dispositions, as well as sensitivity to defeating evidence. Says Greco:

> One aspect of cognitive integration concerns the range of outputs – if the products of a disposition are few and far between, and if they have little relation to other beliefs in the system, then the disposition is less well integrated on that account. Another aspect of cognitive integration is sensitivity to counterevidence, or to defeating evidence. If the beliefs in question are insensitive to reasons that count against them, then this too speaks against cognitive integration (Greco 2010, 152).

Arguably, Temp's disposition to get things right in situations like those described in the *Temp case is* sufficiently integrated into Temp's cognitive character to count as an ability. On the one hand, Temp's cognitive disposition has a fairly wide range of outputs – it is a disposition to

form a wide range of beliefs concerning the temperature in the room and these beliefs concerning room temperature might interact with any number of beliefs formed by Temp's other epistemic dispositions. On the other hand, there is no reason to suppose that the beliefs generated by Temp's disposition are insensitive to counterevidence. We can suppose that Temp might disbelieve that the room temperature is what the thermometer says it is if he has evidence that the thermometer is broken. Given this, it is plausible that Temp's disposition to get things right in cases like the *Temp case* is part of Temp's cognitive character. Temp is the sort of person who will get things right in these sorts of situations; we can contrast Temp again with someone who forms absolutely irrational beliefs on the basis of reading a thermometer – whereas it is part of Temp's cognitive character to form the correct belief regarding room temperature in cases like the Temp case, it is not part of this subject's cognitive character to do so.

It seems, then, that there is no easy way for ability conditions to rule out the *Temp case* as an instance of knowledge. But this should come as no surprise, for well-known reasons. The ability condition's struggle to successfully rule out knowledge in the *Temp case* is just another manifestation of the generality problem for reliabilist epistemologies. Temp's token belief can be said to be the manifestation of many different dispositions, some of which will be reliable, others of which will not be. The difficulty is giving a principled reason for saying why a certain disposition *is* or *is not* relevant. Pure ability accounts of knowledge attempt to solve the problem by claiming that the relevant dispositions are *abilities*. But absent an account of cognitive abilities this only pushes the problem back. And as the remarks above show, it is difficult to give an account of cognitive ability that could be put to use in deciding cases like the *Temp case*. A more promising tack is to give a contextualist account of cognitive ability, and this is exactly what Greco does in more recent work. Says Greco:

> Knowledge-relevant abilities have the following structure.
>
> S has a knowledge-relevant ability A(R/C/D) relative to an environment E = S has a disposition to believe truths in range R when in circumstances C and environment E, with degree of reliability D.
>
> How do we fill in the relevant parameters? For example, what counts as the relevant range of truths, or the relevant sort of circumstances? […] I want to propose […] that we look to relevant information needs to set our parameters. That is, the parameters are set by the needs for information and information sharing attaching to relevant practical tasks (Greco 2012, 17-18).

Greco then goes on to note that the relevant information needs can be either those of the epistemic subjects or those of the people who would attribute knowledge to the epistemic subject (Greco 2012, 19). Greco's account of knowledge-relevant ability, then, is broadly contextualist: there is no simple fact of the matter whether some cognitive disposition is a knowledge-relevant ability; rather a given disposition will be a knowledge-relevant ability relative to the informational needs of epistemic subjects and epistemic attributors.

While this broad contextualism allows for the ability condition to successfully rule that Temp fails to have knowledge, a similar kind of account that replaces talk of *knowledge-relevant abilities* with *safety* can also rule out that Temp has knowledge. For given our informational needs it is natural to describe Temp's method of belief formation as forming a belief on the basis of a broken measuring instrument, in which case it follows that he *is* at serious risk of forming a false belief. And again, given our informational needs we should treat possible worlds in which there is no one manipulating the temperature to match the thermometer as nearby when evaluating the *Temp case*, in which case it follows again (on plausible assumptions) that Temp *is* at serious risk of forming a false belief. So while a broadly contextualist construal of the ability condition can rule out the Temp case as an instance of knowledge, so too can a broadly contextualist construal of the safety condition.

Of course, the mixed epistemologist might reject Greco's contextualism for knowledge-relevant abilities. But in this case, she must provide some other account of what makes a

cognitive disposition an ability. Until she does so, both safety and virtue theories are on equal footing when it comes to ruling out cases like the *Temp case*. The upshot of this is that if cases like the *Temp case* are a real problem for safety conditions to deal with, appealing to ability conditions does nothing to help.

4.4 DIAGNOSIS AND CONCLUSION

In the past section I have argued that cases like the *Barney case* and cases like the *Temp case* are equally problematic (or unproblematic) for pure safety and pure ability accounts of knowledge. Why is this so? The answer seems to be that both safety conditions and ability conditions are different ways of spelling out the intuition that knowledge must come from a reliable source in the right way (cf. Goldman (2012) 70, 87-90). But if contextualists like Heller (1995) are right, there is no fact of the matter whether or not a given belief is reliable. Relative to one context of evaluation, Barney's belief is highly reliable; from another context of evaluation, it is not. Again, relative to one context of evaluation, Temp's belief is highly reliable, relative to a different context of evaluation it is not. It seems that it is the relativity of reliability that allows the putative counterexamples to safety conditions and ability conditions to get a grip. The counterexamples are compelling only to the extent that we fail to notice that we shift from one context of evaluation to another when we judge whether a belief is safe, or formed by an ability, and when we judge whether the belief is an instance of knowledge. But if we keep the context of evaluation fixed by starting out with a judgment that an individual lacks knowledge in a given case, we can go on to describe the individual as failing to meet both the ability and safety condition. And likewise, if we keep the context of evaluation fixed by starting out with a

judgment that an individual possesses knowledge in a given case, we can go on to describe the individual as meeting both the ability and safety conditions on knowledge. My point is that if the counterexamples to both pure safety and pure ability accounts of luck arise because of the context sensitivity of reliability attributions, which is to say here, the context sensitivity of safety and ability attributions, merely combining safety and ability conditions will be no help in dealing with these counterexamples.

In this paper I have argued against mixed accounts of knowledge. The main motivation for mixed accounts is that they are supposed to easily deal with cases problematic for their pure rivals. If I am right, this motivation is illusory. Indeed, it must be if both safety conditions and ability conditions are ways of expressing reliability conditions on knowledge.

One possible response to my argument is to reject both safety accounts, ability accounts and mixed accounts of knowledge, and seek some other way of dealing with these problematic cases. Another possibility is to question whether these cases are problematic after all. There is some reason for pursuing this line; in recent years more and more epistemologists have expressed doubts as to whether knowledge is lacking in so-called dangerous Gettier cases; and I have given reason for supposing that both pure safety accounts and pure ability accounts can deal with cases like the *Temp case* above. Given this, we might take a no-priority view between safety accounts and ability accounts of knowledge. On this view, meeting the safety conditions on knowledge entails that one meets the ability conditions on knowledge and vice-versa. As a result neither safety nor ability conditions are seen as having priority over the other, but both are seen as different ways of spelling out a fundamental intuition concerning knowledge.

4.5 WORKS CITED

Bogardus, T. (2014) "Knowledge Under Threat" *Philosophy and Phenomenological Research* 88:289-313.

Comesaňa, J. (2005). "Unsafe Knowledge" *Synthese* 146:395-404.

Goldman, A. (2012). "Reliabilism" in *Reliabilism and Contemporary Epistemology* (Oxford: Oxford University Press) pp. 68-94.

Greco, J. (2012). "A (Different) Virtue Epistemology" *Philosophy and Phenomenological Research* 85: 1-26.

----. (2010). *Achieving Knowledge* (Cambridge: Cambridge University Press).

Heller, M. (1995). "The Simple Solution to the Generality Problem" *Nous* 29:501-515.

Kelp. C. (2013). "Knowledge: The Safe Apt View" *The Australasian Journal of Philosophy* 91:265-278.

Lackey, J. (2009). "Knowledge and Credit" *Philosophical Studies* 142:27-42.

Lycan, W. (2006) "On the Gettier Problem Problem" in *Epistemology Futures* ed. S. Hetherington (Oxford: Oxford University Press) pp. 148-168

Millikan, R. (1995). "Naturalist Reflections on Knowledge" in *White Queen Psychology and Other Essays for Alice,* (Cambridge: The MIT Press) 241-264.

Pritchard, D. (2007). "Anti-Luck Epistemology" *Synthese* 158:277-297.

----. (2012) "Anti-Luck Virtue Epistemology" *The Journal of Philosophy* 109: 247-279.

----. (2005). *Epistemic Luck* (Oxford: Oxford University Press).

----. (2009). "Safety-Based Epistemology: Whither Now?" *The Journal of Philosophical Research* 33-45.

Schafer, K. (2014). "Knowledge and Two-Forms of Accidental Truth" *Philosophy and*

Phenomenological Research 89: 373-393.

Sosa, E. (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Vol. 1* (Oxford: Oxford University Press).

----. (2010). "How Competence Matters in Epistemology" *Philosophical Perspectives* 24:465-475.

Szabo-Gendler, T. and Hawthorne, J. (2005). "The Real Guide to Fake Barns: A Catalogue of Gifts for your Epistemic Enemies" *Philosophical Studies* 124: 331-352

Turri, J. (2011). "Manifest Failure: The Gettier Problem Solved" *The Philosophers Imprint* 8:1-11.

----. (2015). "From Virtue Epistemology to Abilism: Theoretical and Empirical Developments", in *The Character Project* ed. C. Miller (Oxford: Oxford University Press).

5.0 TRUTH CONNECTEDNESS, GENERICS, AND THE GETTIER PROBLEM

5.1 INTRODUCTION

Most philosophers agree that subjects in Gettier cases fail to have knowledge (though see Hetherington (1999), (2012), Weatherson (2003) and more recently Olsson (2015) for dissent). And most philosophers would agree that subjects in Gettier cases fail to have knowledge because the subjects' grounds[20] for belief and the truth of their beliefs are not connected in the right way. We can call the way that grounds for belief relate to the truth of one's belief, the truth-connectedness of the belief. Most philosophers, then, believe that proper truth-connectedness is necessary for a belief to be an instance of knowledge and that subjects in Gettier cases fail to know because their beliefs do not have proper truth-connectedness. But there is little consensus beyond this. In particular philosophers disagree as to how we should characterize proper truth-connectedness.

In this paper I want to articulate and defend a novel account of proper truth-connectedness. On my account, an individual S's grounds for believing that p are connected in the right way with the truth of p just in case S exemplifies a generic truth that connects having her grounds for belief with believing truly.

[20] 'Grounds' should be taken in a broad sense, so that not only will evidence for p or reasons for believing that p count as one's grounds, but so will using a certain method of belief formation or belief forming process.

I motivate this account by (i) showing how it avoids the major problems of its rivals, (ii) giving reason for thinking that it is extensionally adequate, and (iii) arguing that it best explains why Gettierized beliefs fall short of knowledge.

The paper, then, will proceed as follows. In the first section, I will give a brief review of the two main kinds of accounts philosophers have given of proper truth-connectedness, namely modal and causal/explanatory accounts. I will briefly review some of the general problems facing these accounts. In the second section, I will then lay out my own account of proper truth-connectedness. Since my account appeals to exemplifying generics, we can call it the generic account of proper truth-connectedness. In the third section, I will show that my generic account does not face the problems that modal or causal/explanatory accounts face, and that it is extensionally adequate. Finally in the fourth section I will show that my generic account best explains amongst its rivals why Gettierized beliefs fall short of knowledge.

5.2 EXISTING ACCOUNTS OF PROPER TRUTH-CONNECTEDNESS.

Most philosophers agree that proper truth-connectedness is necessary for knowledge. And most philosophers agree that subjects in Gettier cases fail to have knowledge because their true beliefs lack proper truth-connectedness. Indeed, it seems that this is the best way of characterizing Gettier cases: Gettier cases just are cases in which one fails to have knowledge, not because one's grounds for belief are insufficient for having knowledge, but because one's grounds for belief are not properly connected to the truth of one's belief.

Beyond this, however, there is a lack of consensus. Nevertheless, we can distinguish between three broad approaches to proper truth-connectedness that have carried the field. On the one hand, there are modal accounts of proper truth-connectedness. Nozick (1981), for instance, holds that we can rule out Gettier cases by appealing to a sensitivity condition on knowledge. According to the sensitivity condition on knowledge, an individual S knows that p (using a method of belief formation m), only if S would not believe that p if p were not true. Other philosophers (most notably Pritchard (2005, 2015) have appealed to the safety condition to rule out Gettier cases. Put roughly, the safety condition holds that S knows that p only if in all nearby worlds in which S forms the belief that p on the same grounds she uses in the actual world, S believes truly that p.

Other philosophers hold that we should understand proper truth-connectedness in causal or explanatory terms. So Goldman (1967) famously argued that an individual S knows that p only if there is a causal connection between p and S's believing that p. More recently, Bernecker (2011) has argued for what he calls identificationism. According to Bernecker, in order for S to know that p, the truth-maker of p must either be identical or causally related to the state of affairs that one's justification for believing that p is grounded in. Likewise, Goldberg (Forthcoming) has argued that an individual's belief is subject to epistemic luck just in case her coming to form the belief in question using the method she used is not explainable in terms of the propositions to which she is entitled. Virtue epistemologists such as Greco (2012), Mirachi (2015), Sosa (2007, 2011), and Turri (2011, 2015) have argued that we should understand the right connection in

explanatory terms. These philosophers would agree that an individual S knows that p only if S has the *true belief* that p because it was produced by S's epistemic competence.[21]

Other philosophers hold hybrid views: i.e., views that appeal to modal and explanatory/causal conditions. So for instance, Murphy and Black (2012) combine the sensitivity condition with an explanatory condition on knowledge (cf. Setiya 2012 ch. 3 for a similar approach to accidental true belief).

The purpose of this paper is not to criticize such accounts but to offer an alternative to them. However, I should mention at the outset two general difficulties these accounts face. On the one hand, modal accounts of proper truth-connectedness face problems related to what Shope (1978) calls the conditional-fallacy, but what I think might better be called the problem of *accidentally true counterfactuals*. This is a general problem for counterfactual or modal accounts regardless of their subject matter. The problem is that there are many states of affairs that can make a counterfactual conditional true, and many that could make it false, so that rarely (if ever) can having a substantive property be understood merely in counterfactual terms. To take an example not pertaining to the notion of truth-connectedness, consider the simple conditional account of dispositions. According to the simple conditional account of dispositions, an individual S has a disposition to X in circumstances C just in case S would X if it were in C. (We can put this in terms of possible worlds by saying that an individual S has a disposition to X in conditions C just in case at the nearest possible world in which C obtains S Xs.[22]) Martin

[21] Though amongst these philosophers there is disagreement about what the relevant epistemic competences are, and how exactly we should understand the relevant because. Mirachi as opposed to the others holds that we can understand the relevant competencies only in terms of knowledge. Greco thinks that we can understand the 'because' in causal terms. Sosa and Turri on the other hand think of the relevant sense of 'because' in terms of manifesting a disposition.

[22] Or, if you prefer, at the majority of nearby possible worlds in which C obtains S Xs.

(1994) and Bird (1998), however, give two sorts of examples in which the conditional analysis fails. Martin gives an example in which it is stipulated that an individual (a wire) has a given disposition (it is disposed to shock anyone who touches it) but it is further stipulated that an electro-fink will cut the current if anyone touches it. In this case, it seems that the conditional [the wire is disposed to shock anyone that touches it] is false, but not because the wire lacks this disposition. Rather it is false because if some touches the wire the electro-fink interferes and takes away the disposition. Bird, on the other hand, has us consider a situation in which the following counterfactual conditional is true: if anyone were to consume arsenic they would immediately be given an antidote, so that the arsenic would not harm them. In this case, while it is still true that arsenic has a disposition to harm those who ingest it, the counterfactual conditional, [arsenic would harm anyone, if they were to ingest it] is false.

Similar sorts of counterexamples have popped up for modal accounts of proper truth-connectedness. So for instance, Hiller and Neta (2007) give the following case:

> Orenda believes on good evidence that Jones owns a Ford, and forms the belief that Jones owns a Ford or Brown is in Barcelona. As in the original Gettier case, the belief is justified and true but Orenda does not have knowledge: Jones does not own a Ford but Brown is indeed in Barcelona, unbeknownst to Orenda. But let's also assume that it is no mere accident that Brown is in Barcelona: Brown is so constituted (psychologically, financially, and otherwise) that he would only leave Barcelona in the strangest of circumstances.... In nearly all if not all nearby possible worlds, Orenda's belief is true, but she does not have knowledge (Hiller and Neta 2007, 308).

Hiller and Neta's Orenda case, of course, is an analogue of one of Gettier's original examples, *the Jones owns a Ford or Brown is in Barcelona case*. But their case is tweaked so that, while the truth of Orenda's belief is not properly connected to the grounds of her belief, Orenda's belief meets the safety condition for knowledge. A number of similar cases have been produced in the literature (cf. Stone (2013) and Goldberg (forthcoming) for the most recent examples).

Causal and explanatory accounts of proper truth-connectedness, on the other hand, run into the problem of *deviant causal chains.* An account of some phenomena A will be a causal account just in case it has the following form:

Causal Account: A if and only if C causes E.[23]

Closely related to *causal accounts* are *explanatory accounts*. Explanatory accounts will have the following form:

Explanatory Account: A if and only if C explains E.

Whereas modal accounts face problems related to accidentally true counterfactuals, causal or explanatory accounts face the problem of deviant causal or explanatory chains. Again this is a general problem that faces causal/explanatory accounts regardless of their subject matter. The problem is that there are many ways some object or event might cause an event to occur. As a result it is rarely (if ever) possible to give sufficient conditions for possessing a property merely in causal terms.

To take an example not pertaining to truth-connectedness, consider simple causal or explanatory accounts of intentional action:

> **Intentional Action Causal**: S φs intentionally if and only if S's φing is caused by her intention to φ.
>
> **Intentional Action Explanatory**: S φs intentionally if and only if S's φ's because she has an intention to φ.

These simple sorts of accounts are known to face serious counterexamples. Consider for instance Davidson's (1980/2001) famous mountain climber case:

> **Nervous Mountain Climber:** A mountain climber forms the intention to let go of his partner; but having the intention to let go of his partner causes the climber to become so nervous he inadvertently loosens his grip, letting go of his partner. Intuitively, the mountain climber's letting go of his partner is not intentional in

[23] I take this characterization from Stout (2005, 86).

> this case. However, it is caused by the intention to let go of his partner. And the mountain climber lets go of his partner because he has the intention of letting his partner go.

Cases like Davidson's *Nervous Mountain Climber* show that simple causal or explanatory accounts of intentional action cannot work. For while the nervous mountain climber's letting go of his partner is explained or caused by his having the intention to let go of his partner it isn't explained or caused by the intention in the right way.

Similar sorts of cases have been constructed for causal/explanatory accounts of proper truth-connectedness. My favorite is from Turri (2011).

> (HOBBLED) A competent, though not [masterful] inspection of the crime scene would yield the conclusion that a man with a limp murdered Miss Woodbury. Holmes saw through it and had already deduced that Dr. Hubble poisoned the victim under pretense of treating her.
>
> Holmes also recognized that the scene would fool Watson, whose own inspection of the scene was proceeding admirably competently, though not masterfully... 'Look at him,' Holmes thought, 'measuring the distance between footprints, noting their comparative depth, and a half dozen other things, just as he ought to. There's no doubt where this will lead him -- think how discouraged he will be.' Holmes then resolved, 'Because he's proceeding so competently, I'll see to it he gets it right!'
>
> Holmes sprang into action. Leaving Watson, he hastily disguised himself as a porter, strode across the street to where Hubble was, and kicked him so hard that Hubble was thereafter permanently hobbled with a limp (Turri 2011, 5).[24]

In Turri's *Hobbled* case we see an agent whose belief is produced by cognitive ability/epistemic virtue/competence and is true because it is produced by such an ability. Still, it isn't caused in the right way. As a result it still seems that the connection between Watson's grounds for believing and the truth of his belief is improper.

[24] For a more literary example check out Borges's short story "Death and the Compass" (*La Muerte y la Brujala)*.

Finally, it seems that hybrid accounts don't help matters. This is because the problem of accidentally true counterfactuals and causal/explanatory deviancy can be combined. Consider the following case.

Orenda's *Bénin Genie*

> Orenda is lucky. You've heard of Descartes's famous *malin genie*. Well, she has a *bénin genie*. The bénin genie steps in when Orenda is about to form a false belief concerning a present or future contingent, the bénin genie keeps her from doing so. But instead of changing Orenda's belief, the bénin genie changes the world so that her belief is true. However, Orenda's bénin genie wants her to develop a good cognitive character. Because of this, it intervenes in the world only when Orenda is about to form a false belief but is exercising a cognitive competence. Now it turns out that Orenda has been deceived by Jones into believing that Jones own a Ford, through no fault of her own. On this basis, Orenda is about to form the belief that either Jones own a Ford or Brown is in Barcelona. She is exercising a cognitive competence, but she unfortunately is about to be led into error as a result. The bénin genie steps into action and transports Brown to Barcelona. Furthermore, the bénin genie would always do so, so that in all of the nearest possible worlds in which Orenda forms her belief in the same way, the bénin genie would have made sure Brown was in Barcelona.

In this case we have both problems arising from the conditional fallacy and causal/explanatory deviancy. Hybrid theories, then, will struggle in the same way that modal or causal/explanatory theories do.

The comments above are not meant to be a decisive refutation of modal, explanatory/causal, or hybrid accounts of proper truth-connectedness. Much philosophical work has been put into overcoming the problems that affect these accounts, and while it is controversial, many philosophers believe that they have overcome these problems. I cannot give an overview of all of the proposed solutions here. Nevertheless, the problems faced by these sorts of accounts can motivate the search for another general kind of account, and that is what I offer in the next section.

5.3 THE GENERIC ACCOUNT

In this section I want to articulate what I call the generic account of proper truth-connectedness. Here is the account I favor:

> **Generic:** S's grounds for believing that p are connected in the right way with the truth of p just in case S exemplifies a generic truth that connects having her grounds for belief with believing truly.

To understand this account, we need to get a grip on the notion of exemplifying a generic truth. And to do this we need to get a grip on the notion of a classifying generic proposition.

Classifying generics are statements that make general claims regarding a kind. These statements can have either a bare plural, or a definite or indefinite singular noun phrase. For instance, all of the following sentences can be read as being classifying generics expressing the same proposition:

> Humans have 32 teeth.
>
> The human has 32 teeth.
>
> A human has 32 teeth.[25]

Classifying generic propositions have unique truth conditions. Classifying generics do not simply express universal or statistical generalizations. Nor do they express existential propositions. They do not express universal generalizations, because they can admit of exceptions. (Some human beings don't have 32 teeth, but the generic above is still true.) They do not express statistical generalizations because they can admit of a majority of exceptions. (If tomorrow, for some bizarre reason every human were to have her teeth knocked out, the generic

[25] This example comes from Anscombe (1982). In her paper "Modern Moral Philosophy" she suggests that classifying generics or what she calls *Aristotelian categoricals* are important for doing ethics. These ideas have been further developed by Philippa Foot (2001) and Michael Thompson (2008).

claiming that humans have 32 teeth would still be true.) Finally, they are not existential statements. (Some humans have no teeth, after all, but the generic 'humans have no teeth' is false.)

Some philosophers and linguists hold the 'simple view' regarding the logical form of classifying generics (cf. Carlson (1982), Liebesman (2011)). On this view, a classifying generic is true just in case the kind has the property attributed to it. E.g. The generic, 'tigers are striped' has the following logical form 'Tiger(striped)'. Such a view is true just in case the kind *tiger* is striped.[26] Most philosophers and linguists accept what Leslie-Paul (2015) calls the standard view. According to the standard view the logical structure of classifying generics is tripartite with a non-restricted variable binding operator *Gen* relating sets of conditions with at least one free variable to other sets of conditions. On this view the logical form of 'tigers are striped' is as follows: *Gen* [Tigers(x)][Striped (x)]. My reliance on classifying generics in what follows is compatible with either view, given that we accept a disquotational semantics for *Gen* [Tixers(x)][Striped(x)].

This is because I assume that classifying generic propositions are non-quantificational generalizations. As Leslie (2007) puts things:

> Generics … express generalizations in the sense that they involve commitments as to the nature of previously unencountered instances, but they are not quantificational. They are generalizations that are not fundamentally about *how much* or *how many* (Leslie 2007, 394).

The idea of a non-quantificational generalization can smack of paradox. To understand this claim better, go back to our example of a true classifying generic, "The human being has 32 teeth". To say that this statement expresses a non-quantificational generalization is to say that

[26] This treatment of classifying generics is implicit in the work of Thompson (2008) and Boyle (Forthcoming).

this statement is about human beings in general, attributing to them the property of having 32 teeth, but it is not to attribute to any number of human beings this property.

The idea of non-quantificational generalizations loses some of its paradoxical flavor, if following Prassada and Dillingham (2006) and Prassada (2010), we distinguish between classifying generics that represent a *principled* connection between a kind and some property, and classifying generics that represent *a merely factual* connection between some kind and a property. A classifying generic represents a principled connection between a kind and a property if and only if the following holds: first, the classifying generic can be used to make formal explanations and, second, the classifying generics can ground evaluative inferences. Formal explanations are explanations of why some individual has a given property that appeal to the fact that the individual is a member of some kind. Prassada and Dillingham (2006) give us the following example of formal explanation. When asked, why it is that an individual dog had four legs, people will typically accept the following sort of explanation: this individual dog has four legs because it is a dog, and dogs have four legs. The explanation of why the individual is the way it is works by pointing out a particular classifying generic. Classifying generics that represent merely factual connections cannot be used for formal explanations. So for instance, when asked why a particular barn was red, the majority of people reject the explanation that the barn is red because it is a barn and barns are red. Classifying generics that represent a principled connection between a kind and a property also ground evaluative inferences. So for instance, most people agree that there is something wrong with an individual dog that has three legs, because dogs have four legs. That is, classifying generics that represent a principled connection between members of a kind and having a certain property license the following kind of inference:

Fs are G.

This individual F is not G.

Therefore, this individual F is defective in not being G.[27]

Again, classifying generics that represent merely a factual connection between a kind and property do not ground evaluative inferences. So for instance, while most people will accept the classifying generic dogs are brown, they will not take a non-brown dog to be for this reason defective.[28]

My generic account of proper truth-connectedness appeals to generic propositions, and in particular to generic propositions that represent a principled connection between a kind and having some (complex) property. It assumes that when these propositions are true they represent a sui generis kind of fact.

What motivates this appeal to generic truths? At the beginning of the paper I stated there was broad consensus amongst philosophers that Gettierized beliefs failed to be knowledge because in Gettier cases the subjects' grounds for belief were not properly related to the truth of their beliefs. A popular way of putting this idea is that Gettierized subjects' beliefs are accidentally true or true as a result of luck. This idea can be seen as a motivating both modal and causal/explanatory accounts of proper truth-connectedness. Pritchard, for instance (2005, 2015), explicitly motivates his modal account of proper truth-connectedness by arguing that Gettierized subjects' beliefs are true by luck, and that luck is a modal notion. Likewise, various philosophers (e.g. Greco (2012), Setiya (2012), and Riggs (2014)) have motivated causal/explanatory accounts of proper truth-connectedness on the grounds that accidentality or

[27] Anscombe (1982), Thompson (2008), and Foot (2001) endorse this inference scheme and use it to give an account of natural goodness.

[28] In effect, merely factual classifying generics are vague statistical generalizations.

luck is properly understood in causal/explanatory terms. So for instance, Greco holds that, in general, we can distinguish between activities that are successful because they were the result of an individual's ability or skill and activities that are successful because of luck, i.e., those activities that are not successful because they were the result of an individual's abilities or skill.

I too want to motivate my generic account of proper truth-connectedness by considerations about luck and accidentality. While I don't want to deny that there are senses of accidentality or luck that might best be captured in terms of some modal conditions or in causal explanatory conditions, I believe that there is an important sense of luck or accidentality that is best understood in terms of generic propositions. And I believe that it is this sense of luck that is at play in Gettierized beliefs.

Consider the following account of luck that appeals to generic propositions representing a principled connection.

> **GenLuck**: It is a matter of luck that S Φs (or alternatively, it is an accident that S Φs) just in case S is an F, and there is a true generic proposition representing a principled connection between Fs and Φing but S does not exemplify this true proposition.

What is it to exemplify such a classifying generic? We can contrast an individual S exemplifying the generic 'Ks are F' with an individual S merely being an example of a K that is F. The distinction between being an example of a K that is F and exemplifying the truth that Ks are F is not a distinction without a difference. Consider for example, the true generic *Ravens are black*. Suppose we take a white raven, call him Edgar, and dye him black. In this case Edgar is an example of a raven that is black. But Edgar does not exemplify the true generic *Ravens are*

black.[29] For if ravens were all black in the way that Edgar is black, it seems that *Ravens are black* is false, and *Ravens are white* is true. Or consider the following true classifying generic: The human has 32 teeth. Suppose that Smith has 32 teeth, but only because he originally grew 36 teeth and had four of them knocked out in a hockey fight. In this case, Smith is an example of a human with 32 teeth, but he does not exemplify the generic truth that human beings have 32 teeth. For if everyone had 32 teeth in the way that Smith did, it seems that 'human beings have 32 teeth' would no longer be a true generic statement, but rather 'human beings have 36 teeth' would be true. So an individual S's exemplifying a true generic proposition to the effect that Ks are F goes beyond the simple fact that S is F.

I take the notion of exemplifying a generic truth as basic. But I think in most cases we can give meaningful conditions on when an individual exemplifies a generic truth. Take again a true classifying generic proposition Ks are F. If there are other true generic propositions that explain how it is that Ks are F, then an individual S exemplifies the generic truth, Ks are F, just in case S exemplifies those generic truths that describe how it is that Ks are F. So for instance, it is true that human beings have 32 teeth; but there are also truths concerning how it is that human beings come to have 32 teeth; these are the facts concerning our dental development. Smith then, has 32 teeth, but he fails to exemplify the generic truth that human beings have 32 teeth, because he fails to have 32 teeth in the way that human beings have 32 teeth; his dental development went along another path, and it is only by accident that he as a human has 32 teeth. Or again, take Edgar our dyed black raven. There is a true generic that ravens are black, but there are also true generics that have to do with how it is that ravens are black; how they come to

[29] This example comes from Nickel (2010).

have and maintain their color. And Edgar does not exemplify these truths, and so does not exemplify the generic truth that ravens are black.

Now it is not clear that 'ravens are black', and 'humans have 32 teeth' represent a principled connection between ravens and being black or humans and having 32 teeth. For what it's worth, I believe that they do, but we can find less controversial examples. Take for example the true generic, *dogs have four legs*. Take an individual dog, *Biscuit*. Suppose Biscuit developed in such a way that she originally had five legs, but in a dog fight Biscuit's fifth leg was torn off. In this case there is an important sense in which it is an accident or a matter of luck that Biscuit has four legs, and this is true regardless of how easily or not Biscuit might not have four legs. The idea is that it is an accident or a matter of luck that Biscuit is an example of a dog with four legs.

One of the important features of it being a mere accident that Biscuit has four legs is that the true classifying generic that dogs have four legs cannot be used to offer formal explanations for why Biscuit has four legs; nor can it be used to evaluate Biscuit in terms of the number of legs she has. First, it is not a good explanation for why Biscuit has four legs to point out that Biscuit is a dog and dogs have four legs. It simply isn't true that Biscuit has four legs for this reason. Likewise, it isn't true that Biscuit, in having four legs is not a defective specimen of a dog. Biscuit is a defective specimen (though in some sense the defect of having too many legs has been removed).

My account of proper truth-connectedness can be understood as an anti-luck condition on knowledge, but one that is put forward in terms of *generic* luck. On my account S's belief that p is subject to epistemic luck just in case in believing that p she believes truly on the basis of such and such grounds, but does not exemplify a true generic that connects having these grounds with

believing truly. So for instance an individual forms a true belief on the basis of sense experience, but she does not exemplify a true generic that connects believing truly with basing one's belief on sense experience. On the other hand, S's belief that p is not subject to epistemic luck just in case in believing that p she believes truly on the basis of such and such grounds and *and* she exemplifies a true generic that connects believing truly with basing one's belief on sense experience.

Now clearly, my account of proper truth-connectedness assumes that there are true generic propositions that connect epistemic subjects having certain kinds of grounds for belief and believing truly. Is this assumption defensible? I believe it is. Suppose for instance, that the relevant kind we belong to is *human being*. Are there true generic propositions that connect being a human being with having grounds of a certain kind and believing truly? It seems that there are.

Suppose that you were to meet an extra-terrestrial ethnographer who was interested in how human beings formed true beliefs. The extra-terrestrial ethnographer asks you how it is that human beings form true beliefs. What would you tell him? You would likely say the following sorts of things.

> **Perceiving Is Believing:** Human beings form true beliefs that some state of affairs obtains on the basis of having an experience as though the state of affairs obtained.
>
> **Told You So:** Human beings form true beliefs on the basis of being told by others that some state of affairs obtains.
>
> **Remember the Days:** Human beings form true beliefs about the past on the basis of remembering what took place.
>
> **Jumping Off:** Human beings form true beliefs on the basis of inductive and abductive inference.
>
> **Cluedo:** Human beings form true beliefs by deducing them from known truths.

These classifying generics describe the various ways that human beings come to truth. Mirroring the language of Goldberg (Forthcoming) we can call these ways of coming to truth, routes to truth. The descriptions of these routes to truth all connect being a human being with believing truly on the basis of possessing certain kinds of grounds in a principled way. Consider a human being who does not form true beliefs on the basis of perception, or testimony or memory, or inductive, abductive, or deductive inference. Such an individual would have serious cognitive defects.

The list of true classifying generics I have given above is not supposed to be exhaustive. The thought, however, is that whenever there is a general way that human beings possess truth there will be a corresponding generic connecting grounds for belief (i.e., the relevant method of belief formation) with believing truly.

We can apply this non-reductive account of exemplifying a generic proposition to the generic propositions above that connect being a human being, having certain kinds of grounds for belief, and believing truly. Take for instance *Seeing Is Believing***:** Human beings form true beliefs that some state of affairs obtains on the basis of *seeing* that the state of affairs in question obtains. It seems that there are a range of true classifying generics that we can appeal to that can explain how it is that humans form true beliefs by seeing that something is the case. The details of this description are controversial, as controversial as the philosophy of perception. But I take it that there are clear cases in which one does not exemplify any generic statement describing how human beings form true beliefs by seeing that something is the case. Take for example the following Gettierized belief:

> **Chisholm's Sheep:** Smith sees what looks to be a sheep in the field in front of him. On the basis of this experience he forms the belief that there is a sheep in the field in front of him. As it turns out, what Smith is looking at is actually a cleverly disguised dog. It turns out, however, that there is a sheep in the field out of Smith's line of vision.

In *Chisholm's Sheep* case, Smith is a human being who forms a true belief with such and such content, on the basis of his perceptual experience, where this perceptual experience is good grounds for forming a belief with the relevant content. But Smith does not exemplify *Seeing is Believing*. While it is controversial how it is exactly that humans form true beliefs that some state of affairs obtains on the basis of having an experience as though the state of affairs obtains, it is not controversial that Smith isn't forming a true belief on the basis of experience this way.

Mutatis mutandis the same holds for other Gettier cases. Take for example a modified version of one of Gettier's cases:

> **The New Boss Is Just Like The Old Boss**: Smith is told by Jones, 'The boss has ten coins in his pocket.' Smith knows Jones to be a trustworthy informant, and while he is confused by why Jones is interested in how many coins the boss has in his pocket, he forms the belief that the boss has ten coins in his pocket. It turns out that Brown, the old boss, had just been fired, and company rules stipulate that, Smith is automatically made the boss. Smith happens to have ten coins in his pocket.

In this case, Smith is an example of someone who forms a true belief with such and such content on the basis of being told by someone that such and such states of affairs obtain. But Smith does not exemplify *Told You So.* Whatever the proper generic description of how it is that human beings form true beliefs on the basis of being told by others that some state of affairs obtains, Smith forms his true belief in some different way.

Again we can turn to Gettier cases involving inferential knowledge, and similar remarks apply. Take for instance, Gettier's classic case built around disjunction introduction:

> **Nogot:** Smith takes Jones to be a reliable informant. Jones tells Smith that that Jones has recently bought a Ford and shows Smith what looks to be a title for a new car. Smith, on a whim and knowing absolutely nothing concerning the whereabouts of Brown, forms the belief that Jones owns a Ford or Brown is on Barcelona, on the basis of disjunction introduction and the premise that Jones owns a Ford. As it turns out, Jones has been lying, and does not own a Ford, but Brown is in Barcelona.

In this case Smith is an example of someone who has formed a true belief on the basis of disjunction introduction. Still, Smith clearly does not exemplify *Cluedo*. For whatever the true classifying generics are that describe how human beings form true beliefs on the basis of deductive inference, it is clear that Smith does not form his true belief in these ways.

How does the generic account deal with Red Barn Façade and other similar unpossessed defeater cases? Consider the following case:

> **Barney**: Barney is driving, unawares, through Red Barn Façade County (RBFC). There is only one real red barn in RBFC but hundreds of remarkably realistic red barn facades, barn facades so realistic that they would trick anyone who wasn't in the know into believing that they were real red barns. Barney drives past the only red barn in the county, and forms the belief that there is a red barn in front of him on the basis of his experience of what seems to be (and is) a red barn located directly in front of him.

For a long time, philosophical orthodoxy held that Barney's belief was Gettierized, and that he failed to have knowledge. In recent years however, more and more philosophers have come to hold that Barney does have knowledge and that his belief is not Gettierized. Many of these philosophers' conversions have been motivated by theoretical concerns (cf. Sosa (2007)); others have noted how hard it is to give a principled way of distinguishing red barn façade cases from other cases in which we intuitively do have knowledge (cf. Szabo and Hawthorne (2004)). But

while it is no longer appropriate to speak of 'philosophical orthodoxy' on this matter, there are many philosophers that hold Barney fails to have knowledge because his belief is Gettierized.

Given this state of affairs, an account of proper truth-connectedness should do one of the following: (i) give an account of epistemic luck that implies a clear verdict on red barn façade and other cases, along with an error theory explaining why many philosophers have made false judgments regarding this case; or (ii) give an account of epistemic luck that does not imply a clear verdict on red barn façade cases, but instead explains why such cases are controversial. The generic account I have given above does the latter. Assuming that the generic account of proper truth-connectedness is correct, the question of whether or not Barney's belief is Gettierized comes down to the following question: does Barney's belief that there is a red barn façade in front of him exemplify *Seeing Is Believing?* It is clear that Barney is an example of someone who forms a true belief that things are thus and so on the basis of his experience that things are thus and so. But is Barney an exemplification of someone who does this? All of this depends how it is, generically speaking, that human beings move from experiences of things being thus and so to believing that things are thus and so. *Ex hypothesi*, what is internal to Barney, that is, that part of Barney's movement from experience to true belief that pertains to Barney's cognitive activity, does not prevent Barney from exemplifying *Seeing is Believing*. The question is whether external, environmental factors can keep one from exemplifying *seeing is believing*, and in particular, whether being in an environment in which one could very easily form a false belief on the basis of one's experience can keep one from exemplifying the relevant generic proposition.

There are reasons for and against accepting this proposition. On the one hand, 'normal' conditions are presupposed by classifying generics. Setiya (2012) puts this point well in arguing

that there are not classifying generics (at least of the relevant kind) for *rational beings*. Says Setiya:

> When Fs are by nature G, in the generic sense, though some Fs are not G, that a particular F is G depends on the circumstances in which it finds itself. The conditions required for an F to become G, and in which such development fails, themselves belong to the natural history of the F. They specify the needs that must be met, the internal conditions, and the proper environment, in which an F is G (Setiya 2012, 212-213).

It is reasonable to assume that the normal conditions in which human beings form true beliefs on the basis of experience are such that one is not in danger of forming false beliefs on the basis of experience.

At the same time, however, when we are thinking about whether or not an individual exemplifies a generic proposition, it is unclear what constitutes the relevant environment of the individual. Consider the following toy example. Human beings have 32 teeth. Suppose furthermore, that part of the story of how it is that human beings come to have and keep 32 teeth is by staying in an oxygen rich environment. Now suppose a terrarium is built on the moon, one that is capable of supporting human life. Now, take someone who grows up in the terrarium and who develops 32 teeth. It seems that we can say that such a person exemplifies the generic truth that human beings have 32 teeth, and this is because she exemplifies the ways in which human beings have 32 teeth. But if we consider her environment to go beyond the confines of the terrarium then it seems that she does not exemplify the generic *human beings have 32 teeth*, since this environment on the whole is not an oxygen rich environment.

In the terrarium case, it seems clear that we should only consider the individual's immediate environment inside the terrarium. But things are more complicated with the Barney example. Suppose we take Barney's relevant environment to be his immediate environment, i.e., encompassing everything that is in his perceptual field. In this case, *ex hypothesi*, Barney's

environment is normal, and so there is no reason for supposing that Barney does not exemplify *Seeing is Believing*.

In this paper, I am not interested in arguing how we should think of one's environment in the Barney case. Instead, I want to show how it is that the generic account of proper truth-connectedness can explain why philosophers' intuitions differ regarding Red Barn Façade and other unpossessed defeater cases. One way of explaining the difference in intuitions is pointing to the vagueness of the concept 'environment'. This vagueness allows for differing judgments regarding whether an individual is in a normal environment, and so it allows for differing judgments as to whether an individual has exemplified the relevant generic proposition.

It should be clear that the generic account of proper truth-connectedness easily deals with those cases that were problematic for modal and causal explanatory accounts of proper truth-connectedness. Take the Orenda case again:

> Orenda believes on good evidence that Jones owns a Ford, and forms the belief that Jones owns a Ford or Brown is in Barcelona. As in the original Gettier case, the belief is justified and true but Orenda does not have knowledge: Jones does not own a Ford but Brown is indeed in Barcelona, unbeknownst to Orenda. But let's also assume that it is no mere accident that Brown is in Barcelona: Brown is so constituted (psychologically, financially, and otherwise) that he would only leave Barcelona in the strangest of circumstances.... In nearly all if not all nearby possible worlds, Orenda's belief is true, but she does not have knowledge (Hiller and Neta 2007, 308).

Prima facie, the Orenda case is problematic for safety accounts of proper truth-connectedness. But it is not problematic for my generic account. For just as Smith in the *Nogot* case fails to exemplify *Cluedo*, so too does Orenda. However it is that human beings, generically speaking, form true beliefs on the basis of deductive inference, Orenda is not forming her true belief like that.

Likewise consider the *Hobbled case* again.

> (HOBBLED) A competent, though not [masterful] inspection of the crime scene would yield the conclusion that a man with a limp murdered Miss Woodbury. Holmes saw through it and had already deduced that Dr. Hubble poisoned the victim under pretense of treating her.
>
> Holmes also recognized that the scene would fool Watson, whose own inspection of the scene was proceeding admirably competently, though not masterfully... 'Look at him,' Holmes thought, 'measuring the distance between footprints, noting their comparative depth, and a half dozen other things, just as he ought to. There's no doubt where this will lead him -- think how discouraged he will be.' Holmes then resolved, 'Because he's proceeding so competently, I'll see to it he gets it right!'
>
> Holmes sprang into action. Leaving Watson, he hastily disguised himself as a porter, strode across the street to where Hubble was, and kicked him so hard that Hubble was thereafter permanently hobbled with a limp (Turri 2011, 5).

Again, while *prima facie* the *Hobbled case* is problematic for causal/explanatory accounts of proper truth-connectedness, it is not problematic for the generic account I defend. However it is that human beings, generically speaking, form true beliefs on the basis of inductive inference, Watson is not forming his true belief like that.

5.4 FURTHER ADVANTAGES OF THE GENERIC ACCOUNT

In the last section, I explained the generic account of proper truth-connectedness, and I gave reasons for believing that it is extensionally adequate: it gives the right verdict in standard Gettier cases, explains the differences in judgment of philosophers towards cases like the Red Barn Façade case, and it deals easily with those cases that are problematic for modal or causal/explanatory accounts of proper truth-connectedness. In this final section, I argue that it has a further advantage as well, namely, that it better explains why Gettierized beliefs are not instances of knowledge.

As I noted in the introduction, while the majority of philosophers (and for that matter, the majority of laymen) believe that we fail to have knowledge in Gettier cases, there are important

voices of dissent in the literature. If an account of proper truth-connectedness carries with it an explanation of why Gettierized beliefs are not instances of knowledge, then *ceteris paribus* it is preferable to its rivals.

Why is it that Gettierized beliefs are not instances of knowledge? A satisfying answer to this question will not only point to some quality of Gettierized beliefs, but it will give some indication as to why this quality is incompatible with knowledge. The best answers to the question will be answers that flow from the nature of knowledge. Now, on my view, the core concept of knowledge is normative; and the reality of knowing is an irreducibly normative reality. Wilfrid Sellars in *Empiricism and Philosophy of Mind* puts the idea powerfully as follows:

> [T]he idea that epistemic facts can be analyzed without remainder — even in principle — into non-epistemic facts, whether phenomenal or behavioral, public or private, with no matter how lavish a sprinkling of subjunctives and hypotheticals, is, I believe, a radical mistake — a mistake of a piece with the so-called "naturalistic fallacy" in ethics (§5).

And further on he states:

> [I]n characterizing an episode or a state as that of *knowing*, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says (§36).

We can accept the Sellarsian insight that knowledge is essentially normative without weighing in on the merits of internalist or externalist accounts of knowledge. But given the Sellarsian insight, accounts of knowledge couched in non-normative language conceal what is essential to knowledge. This is true of non-reductive as well as reductive accounts of knowledge. Whether one says that knowledge is non-accidentally true belief or one says that knowledge is the most general stative-factive attitude one leaves it a mystery as to why

knowledge is epistemically superior to merely true belief and one leaves it open that knowledge has no intrinsic normative status.

For this reason, my preferred account of knowledge is that knowledge just *is* non-defective belief. This claim needn't be controversial. Most importantly, it does not require a *reductive* reading. Perhaps to understand non-defective belief we must make appeals to knowledge. Furthermore, there is no reason to doubt the extensional adequacy of this claim if we allow that it is a non-reductive, circular account of knowledge. Of course, one could complain that since it is circular, it is also uninformative. But I doubt that it is. Suppose that we cannot get a grip on non-defective belief unless we already have a grip on knowledge, and vice versa; still, the claim that knowledge is non-defective belief allows us to place knowledge claims in a particular evaluative context -- an evaluative context in which judgments concerning *defect* are intelligible; and this can shine some light on the reality of knowledge.

If knowledge just is non-defective belief, it follows that knowledge attributions must take place in an evaluative context in which judgments concerning defect are intelligible. But in what evaluative contexts are judgments concerning defect intelligible? Here I do not propose to give a full list of such contexts, but I note that judgments concerning defect *are* intelligible in contexts of evaluation concerning *natural goodness*.

On Foot and Thompson's account of natural goodness, judgments concerning natural goodness are regimented by the following inference scheme (cf. Thompson 2007, 80, and Foot 2001, 30):

> The F has such and such properties.
> This particular F lacks these properties.
> Therefore, this particular F is defective.

On this view, an individual S is naturally good in every way just in case it is in no way defective, i.e., just in case it is an F and it has all of the characteristics that Fs are said to possess. Note that the variable F in the inference scheme above can range over not only kinds of animals but kinds related to kinds of animals, in particular it can range over kinds of parts and kinds of characteristic activities or performances of animals.

To see this, consider the following examples:

The beaver has a strong mandible for chewing down trees.
This particular beaver doesn't.
Therefore, this beaver is defective.

The beaver's mandible is strong enough to chew down trees.
This beaver's mandible is not strong enough to chew down trees.
Therefore, this beaver's mandible is defective.

The beaver builds dams to raise the water level above the entrance to its den.
When this beaver built a dam it did not raise the water level above the entrance to its den.
Therefore, this beaver's activity of building a dam was defective.

While it has not been stressed by theorists working on the notion of natural goodness, failure to *exemplify* the relevant generics even if one possesses the relevant quality also constitutes defect. As I noted above, Biscuit the dog who has four legs because she lost her fifth in a dog fight is naturally defective, even if she now has four legs as she should and not five. And likewise, a beaver might build a dam that raises the water level above the entrance to its den, but only by accident, since in doing so he does not exemplify the generic truth that beavers

build dams to raise the water level above the entrances of their den; the beaver's activity, though it achieves its end, is still naturally defective.

Now it should be clear that my generic account of proper truth-connectedness comfortably fits into the Foot-Thompson account of natural goodness. Again, according to the generic account I give, S's grounds for believing that p are connected in the right way with the truth of p just in case S exemplifies a generic truth that connects having her grounds for belief with believing truly. As such, from a natural goodness standpoint, any belief of S's will be defective that does not have proper truth-connectedness. For in this case, S's belief will fail to exemplify a relevant generic proposition connecting believing on the grounds one has to believing the truth. In this case, Gettierized beliefs, beliefs that do not have proper truth-connectedness, are thereby defective. And as they are defective, they thereby fail to be instances of knowledge.

If what I say above is correct, then the generic account of proper truth-connectedness explains in a clear way why Gettierized beliefs fail to be instances of knowledge. This is not so with modal and causal-explanatory accounts of proper truth-connectedness.

5.5 CONCLUSION

In this paper I have articulated and defended a generic account of proper truth-connectedness. The generic account of proper truth-connectedness is intimately related to an account of what I call *generic luck*. I have motivated this account of proper truth-connectedness by showing how it avoids the problems confronting its rivals, by giving reasons for believing that it is extensionally

adequate, and by showing how it can easily explain why Gettierized beliefs are not instances of knowledge.

As noted in the previous paragraph, my account of proper truth-connectedness is closely related to the notion of *generic luck*. It is my belief that this notion deserves further attention regardless of how well my account of proper truth-connectedness fares. Notions of luck or accidentality show up in problematic ways throughout the philosophical literature, and a generic conception of luck or accidentality has been virtually ignored. I have maintained in this paper that the notion of generic luck as I have articulated it, illuminates epistemic luck and proper truth-connectedness. I am confident that it can illuminate other areas of interest.

5.6 WORKS CITED

Anscombe, G. (1981) "Modern Moral Philosophy" in *Ethics Religion and Politics: Collected Philosophical Paper Vol. III* (Oxford: Wiley-Blackwell).

Bernecker, S. (2011). "Keeping Track of the Gettier Problem" *The Pacific Philosophical Quarterly* 92:127-152.

Bird, A. (1998) "Dispositions and Antidotes" *The Philosophical Quarterly* 48:227-234.

Davidson, D. (1980/2001). *Essays on Action and Events* (Oxford: Oxford University Press).

Foot, Philippa. (2001). *Natural Goodness*. (Oxford: Oxford University Press).

Goldberg, S. (Forthcoming). "Epistemic Entitlement and Luck" *Philosophy and Phenomenological Research*.

Goldman, A. (1967). "A Causal Theory of Knowing" *The Journal of Philosophy* 64:357-372.

Greco, John. (2012). "A (Different) Virtue Epistemology" *Philosophy and Phenomenological*

Research 85: 1-26.

Hetherington, S. (2012). "How to Know: A Practicalist Conception of Knowledge" *Philosophia* 40:539-547.

----. (1999). "Knowing Failably", *The Journal of Philosophy* 96:565-587.

Hiller A. and Neta, R. (2007). "Safety and Epistemic Luck," *Synthese* 158:303-313.

Leslie, S. (2007) "Generics and the Structure of the Mind" *Philosophical Perspectives* 21: 375-403

----. (2008). "Generics: Cognition and Acquisition" *Philosophical Review* 117: 1-47.

----. (2015). "Generics Oversimplified" *Nous* 49:28-54.

Liebesman, D. (2011). "Simple Generics" *Nous* 45: 409-442.

Martin, C. (1994). "Dispositions and Conditionals" *The Philosophical Quarterly* 44:1-8.

Miracchi, L. (2015). "Competence to Know" *Philosophical Studies* 72:29-56.

Murphy, B. and Black, T. (2012). "Sensitivity Meets Explanation: An Improved Counterfactual Account of Knowledge" in *The Sensitivity Principle and Epistemology* Eds. Becker and Black (Cambridge: Cambridge University Press, 28-42).

Nickel, B. (2010). "*Ceteris Paribus* Laws, Generics and Natural Kinds" *The Philosophers Imprint* 10:1-25.

Nozick, R. (1981). *Philosophical Explanations* (Cambridge: Harvard University Press).

Olsson, E. (2015). "Gettier and the Method of Explication: a 60 Year Old Solution to a 50 Year Old Problem" *Philosophical Studies* 172:57-72.

Prasada, S. (2010) "Conceptual Representation and Some Forms of Genericity" in *Kinds, Things and Stuff* ed. Pelletier, (Oxford: Oxford University Press) 36-59.

Prasada S. and Dillingham E. (2006). "Principled and Statistical Connections in Common Sense Perception" *Cognition* 99:73-112.

Pritchard, D. (2015). "Anti-Luck Epistemology and the Gettier Problem" *Philosophical Studies*

172:93-111.

----. (2005) *Epistemic Luck* (Oxford: Oxford University Press).

Riggs, W. (2014). "Luck, Knowledge, and Mere Coincidence" *Metaphilosophy*: 4-5: 627-639.

Sellars, W. *Empiricism and Philosophy of Mind* (Cambridge: Harvard University Press).

Setiya, K. (2012). *Knowing Right from Wrong* (Oxford: Oxford University Press).

Shope, R. (1978). "The Conditional Fallacy" *The Journal of Philosophy* 75:397-413.

Sosa, E. (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Vol. 1*. (Oxford: Oxford University Press).

----. (2011). *Knowing Full Well* (Princeton: Princeton University Press).

Stone, J. (2013). "Unlucky Gettier Cases" *The Pacific Philosophical Quarterly* 94:421-430.

Stout, R. (2005) *Action* (Oxford: Acumen).

Szabo-Gendler, T. and Hawthorne, J. (2005). "The Real Guide to Fake Barns: A Catalogue of Gifts for your Epistemic Enemies" *Philosophical Studies* 124: 331-352.

Thompson, Michael. (2008). *Life and Action: Elementary Structures of Practice and Practical Thought*. (Cambridge, MA: Harvard University Press).

Turri, J. (2015). "From Virtue Epistemology to Abilism: Theoretical and Empirical Developments", in *The Character Project* ed. C. Miller (Oxford: Oxford University Press).

----. (2011). "Manifest Failure: The Gettier Problem Solved." *Philosophers Imprint* 11:1-11

Weatherson, B. (2003). "What Good are Counterexamples?" *Philosophical Studies* 115:1-31.

6.0 GENERICS AND THE GENERALITY PROBLEM

6.1 INTRODUCTION

In this paper I want to articulate and defend a version of reliabilism. I believe that my version of reliabilism successfully deals with the *generality problem*. Like all reliabilists, I hold that there is a conceptual connection between a belief's being justified and a belief's coming from reliable method for belief formation. But my view differs from other reliabilist views in two respects. On the one hand, I understand reliability not in statistical terms, but rather in generic terms. So on my view, a method of belief formation is reliable just in case the following habitual holds true: *This method of belief formation is used by members of my kind to produces true beliefs*. On the other hand, on my view we need to look to an individual's kind in order to determine whether a method of belief formation is relevant for epistemic evaluation. According to my view, the only relevant belief forming methods are those that are connected in *a principled way* to the kind of thing that one is. Furthermore, a method of belief formation is relevant for epistemic evaluation only if one uses the method in a way which has a principled conceptual connection to how members of one's kind use the method. By appealing to habituals, and to principled connections between a kind and various methods of belief formation, my version of reliabilism is able to deal with both aspects the generality problem. Or so I will argue.

This paper will proceed as follows. In the first section, I will lay out a simple version of reliabilism, and show how this sort of reliabilism falls prey to both the generality problem. In the second section I will give my own reliabilist account of justification, and show how it

overcomes the generality problem. Finally in the third, section I will consider some objections to my view.

6.2 THE GENERALITY PROBLEM FOR RELIABILISM

Consider the following simple form of reliabilism.

> **SR**: S's belief that p is justified just in case S forms the belief that p using a reliable method of belief formation, where a method of belief formation is reliable in turn, just in case it produces a high ratio of true beliefs.

No one, to my knowledge, has endorsed a theory as simple as SR, though it does approximate Goldman's view in "What is Justified Belief" (Goldman 1979). Still, it is worth noting that SR has considerable appeal. Its main strength is how easily it accommodates our intuition that there is a conceptual connection between a belief's being justified and a belief's having some truth-conducive properties. On *SR* these truth conducive properties are understood in terms of the provenance of one's belief: a belief is justified just in case it results from a method of belief formation that produces a high ratio of true beliefs.

Reliabilist theories like SR, however, face a serious and well known problem: the generality problem. We can understand the generality problem to have two distinct aspects. Feldman and Conee (1985) set up the first part of the problem as follows:

> Suppose that one evening Jones looks out of his window and sees a bright shining disk-shaped object… Jones is attracted to the idea that extraterrestrials are visiting the Earth. He manages to believe that he is seeing a flying saucer. Is the process that caused this belief reliable? Since the sequence of events leading to his belief is an instance of many types of process, the answer depends upon which of these many types is the relevant one. The sequence falls into highly general categories such as perceptually-based belief formation and visually-based belief formation… The sequence of events leading to Jones's belief also falls into many relatively specific categories such as night-vision-of-a-nearby-object and vision-in-Jones's-precise-environmental-circumstances… The sequence is also an instance of this contrived kind: process-leading-from-obviously-defeated-evidence-to-the-belief-that-one-sees-a-flying-saucer… Finally,

> there is the maximally specific process that occurs only when physiological events occur that are exactly like those that led to Jones's belief that he saw a flying saucer. In all likelihood this kind of process occurred only once. Processes of these types are of differing degrees of reliability, no matter how reliability is determined. (Feldman and Conee 1985, 26).

Feldman and Conee's point is that a simple reliabilist account of justification like SR gives no definite verdict for Jones's belief that there is a flying saucer in front of him, because his belief can be described as being produced by any number of methods of belief formation, some of which are intuitively reliable, some of which are not. To give a clear verdict, the simple reliabilist theory of justification needs to be supplemented with a principled account that determines which types of methods of belief formation are relevant for the epistemic evaluation of Jones's belief, so that Jones's belief is attributed to only one method of belief formation, or at the very least, Jones's belief is attributed only to reliable (or unreliable) methods of belief formation.

The first aspect of the generality problem arises, then, because the same belief can be described as being the product of different methods of belief formation, some of which are reliable and some of which are not. The second aspect of the generality problem, on the other hand, arises because reliability is understood in statistical terms. So for instance, Goldman (1986) defines reliability as follows:

> An object (a process, method, system, or what have you) is reliable if and only if (1) it is a sort of thing that tends to produce beliefs, and (2) the proportion of true beliefs among the beliefs it produces meets some threshold, or criterion, value. Reliability, then, consists in a tendency to produce a high truth ratio of beliefs (Goldman 1986, 26).

For Goldman's definition of reliability to have definite content we need to presuppose a set of beliefs produced by some belief forming process. If the ratio of true beliefs in this set is sufficiently high then the process is reliable. But whether or not the ratio of true beliefs in the set is sufficiently high can depend on what we take the relevant set of beliefs to be. What is to count

as the true beliefs that the process produces? Should we count only the beliefs the process actually has produced? Or should we count all the beliefs that the process might produce? Or only all the beliefs that the process might produce in normal conditions? And should we consider only those beliefs produced by the process in the immediate environment? Or should we consider beliefs produced by the process outside of the immediate environment as well? Depending on how we answer these questions it will turn out that the same belief-forming process is reliable or unreliable. And so a similar difficulty arises like that stressed by Conee and Feldman above: unless simple reliabilism is supplemented with an account of what beliefs make up the relevant reference class, it provides us with no definite verdict regarding even the simplest cases (Cf. Swinburne 2001, 17-19).

While there have been several attempts at solving the generality problem on its own terms (cf. Alston (1995), (Beebe (2004), (Wunderlich 2003), none of the proposals in the literature has won general acceptance. As a result, there seem to be two main sorts of response to the generality problem in the current literature: the first, represented by Comesana (2006) and Bishop (2010) we can call the *tu quoque* response. According to this line of thought, the generality problem is a problem shared by *every* theory of justification and knowledge, and so does not count as reason for rejecting reliabilism. The second approach, represented by Heller (1995) and Greco (2010, 2012) we can call the *contextualist* response. According to this line of response, there is not a principled solution to the generality problem: rather, what types of belief forming methods are relevant for epistemic evaluation, and what beliefs are relevant for determining the reliability of a belief-forming method, is determined by the interests and projects of either the epistemic subject, or those evaluating the epistemic subject's beliefs.

It should be noted that both of these responses are concessive. On the one hand, the *tu quoque* response does nothing to resolve the generality problem; it simply deflects criticism of reliabilist theories by arguing that the generality problem is a problem for everyone. On the other hand, the contextualist response grants that the generality problem cannot be solved in a principled way; it simply deflects criticism of reliabilist theories by arguing that we don't need a principled solution to the generality problem.

6.3 THE GENERALITY PROBLEM FOR RELIABILISM SOLVED

6.3.1 Classifying Generics and The First Aspect of the Generality Problem

In this section I want to begin to put forward a version of reliabilism that can take the generality problem head on. On my view, a belief is justified if it is formed by a *relevant, reliable* belief-forming mechanism in the *right* way; otherwise it is not justified. This is the skeleton of my view. To put flesh on it, I need to clarify what it is for a method of belief formation to be relevant, what is for a belief to be formed by a method in the right way, and what it is for a belief to be reliable. To do so I need to make reference to kinds propositions that have been relatively ignored by philosophers in the analytic tradition. The propositions I have in mind are expressed through *classifying generic* and *habitual* statements.

Classifying generics are statements that make general claims regarding a kind. These statements can have either a bare plural, or a definite or indefinite singular noun phrase. For instance, all of the following sentences can be read as being classifying generics expressing the same proposition:

Humans have 32 teeth.

The human has 32 teeth.

A human has 32 teeth.

The propositions expressed by classifying generics have unique truth conditions. Unlike universal propositions, generic propositions admit of exceptions. *Humans have 32 teeth* is true, even though my daughter only has 5. Indeed, unlike statistical generalizations (*most humans have 32 teeth*) generic propositions of the form Ks are F, can be true even if the majority of Ks are not F. So for instance, for all I know, it may be false that the majority of humans currently living have 32 teeth. Still *humans have 32 teeth* is a true generic proposition. Furthermore, unlike existential propositions, generic propositions of the form Ks are F are not true if some K is F. So while there are human beings with exactly 15 teeth, the classifying generic *humans have 15 teeth* is not true.

Following recent work by Carlson (1982, 2010), Leslie (2007, 2008, 2015) and Liebesman (2011), we can understand classifying generic propositions as non-quantificational generalizations. As Leslie (2007) puts things:

> Generics … express generalizations in the sense that they involve commitments as to the nature of previously unencountered instances, but they are not quantificational. They are generalizations that are not fundamentally about *how much* or *how many* (Leslie 2007, 394).

To understand this claim better, go back to our example of a true classifying generic, "The human being has 32 teeth". To say that this statement expresses a non-quantificational generalization is to say that this statement is about human beings in general, attributing to them the property of having 32 teeth, but it is not to attribute to any number of human beings this property. This might seem odd, if one is accustomed to think of all generalizations to be quantificational – but it is exactly this assumption that theorists like Leslie call into question. Generic statements are a *sui generis* form of generalizations.

We can make the idea of non-quantificational generalizations seem less odd if we focus on the different ways that properties can be predicated of a kind. Following Prassada (2010) we can distinguish between representing a property as having a *principled* vs. a *factual* connection to a kind. According to Prassada (2010) and Prassada and Dillingham (2006) when we represent a property as having a principled connection to a kind, this licenses our making *formal explanations* and *normative judgments* on the basis of this principled connection. A formal explanation represents a property as possessed by individauls of a kind *in virtue* of the individuals being members of this kind. So for instance, Prassada and Dillingham give evidence that people in general agree that dogs have four legs in virtue of being dogs. But most people disagree that dogs are brown in virtue of being dogs. Likewise, most people disagree that barns are red in virtue of being barns. So the true classifying generic *dogs have four legs* represents a principled connection between being a dog and having four-legs. Likewise, Prassada and Dillingham present evidence that people tend to judge individuals of a kind to have a defect when they lack some property which has a principled connection to the kind. So for instance, people tend to judge that there is a defect in a dog that has more or less than four legs because the dog has more or less than four legs. This again contrasts with merely factual connections; non-brown dogs are not judged to be defective on account of their being some other color than brown.

When we focus on the explanatory and normative aspect that mark our judgments concerning principled connections between properties and a kind, it seems less odd to say that classifying generics express non-quantificational generalizations concerning a kind. For we can understand these sorts of generalizations as expressing a principled connection between members

of a kind and a certain property. This principled connection between members of a kind and some property can exist regardless of the number of individuals that possess the property.

So much for classifying generic statements. Let's return to the first aspect of the generality problem for reliabilism. This problem arises because oftentimes the same belief can be thought of as being the product of diverse methods of belief formation, some of which are intuitively reliable, others of which are intuitively unreliable. As a result, in even the simplest cases, reliabilism gives no definite verdict as to whether one's belief is justified. My proposal is that by appealing to classifying generics we can deal with this first aspect of the generality problem. In particular I think that the following claim is true:

> **The Right Kind** A method of belief-formation is *relevant* for the epistemic evaluation of S's belief that p, only if S is a member of kind K, and the following classifying generic is true and expresses a principled connection: Ks use M to form beliefs.[30]

The basic idea is that in order to determine whether a belief forming a method of belief formation is relevant for evaluating an individual's belief, we need to look to the kind to which the individual belongs. It is the methods of belief formation that belong in a principled way to a kind that are relevant for the evaluation of an individual's beliefs.

This idea has intuitive appeal. On the one hand, when we evaluate beliefs, we are not interested in the idiosyncrasies of an individual's psychology. If S's belief is formed by a so-called 'strange and fleeting process' no matter how reliable the process might be, S's belief lacks

[30] Or more fully, a method of belief-formation is relevant for the epistemic evaluation of S's belief that p, only if S is a member of kind K, and the following classifying generic is true and expresses a principled connection: Ks use M to form beliefs in order to arrive at the truth. This addition is to deal with possible methods of belief formation that human beings employ not in order to arrive at the truth, but to achieve some practical end; so for instance, human beings might form beliefs on the basis of wishful thinking, but this method of belief formation is not directed at the truth, but at some practical good, e.g. peace of mind. Setiya (2012) appeals to generic propositions in dealing with the question of non-accidental truth.

justification. Consider, for instance, BonJour's (1985) famous example of a man with who by some cosmic accident has clairvoyant powers, though he has no evidence either for or against his possession of these powers. Intuitively, this clairvoyant lacks justification for those beliefs that he bases on his clairvoyant power alone. Likewise, consider Plantinga's (1993) example of someone with a serendipitous brain lesion that directly causes its possessor to form the true belief that he has a brain lesion. Again, this individual seems to lack justification for his belief, no matter how reliable the process that leads to its formation. *The Right Kind* would explain why this is so in both cases: it does not belong to the human being, or the rational animal to form beliefs in such ways. So these highly reliable ways of forming beliefs that are used in the clairvoyant and brain lesion cases are not relevant for epistemic evaluation.

On the other hand, there does seem to be a list of basic methods of belief formation such that there is a principled connection between employing these methods of belief formation and being a human being, and which, furthermore, are the sorts of methods of belief formation that confer justification upon our beliefs. Consider for instance the following list of classifying generics:

(a) Human beings form beliefs on the basis of testimony.

(b) Human beings form beliefs on the basis of perceptual experience.

(c) Human beings form beliefs on the basis of memory.

(d) Human beings form beliefs on the basis of different sorts of inferences – for instance, deductive inferences, inferences to the best explanation, statistical inferences etc.

It seems that there is a principled connection between employing these methods of belief formation and being a human being. First, we can give formal explanations that connect having and employing these methods of belief formation. It is because that an individual is human that she employs these methods of belief formation in the same way that it is because an individual is

a dog that it has four legs. Second, it seems that we can make normative judgments to the effect that an individual human being is defective if she does not use these sorts of methods of belief formation. Imagine someone who lacks the ability to form beliefs on the basis of deductive inference, or who lacks the ability to form beliefs on the basis of memory; clearly such a person is cognitively defective.

Furthermore, oftentimes one's belief can be justified in virtue of being formed by these methods; one's belief might be justified because it was formed on the basis of perceptual experience, or on the basis of memory, deductive inference etc.

Now, if *The Right Kind* is correct, the only sorts of methods of belief formation that will be relevant in evaluating a subject's beliefs will be those methods of belief formation that have a principled connection with the epistemic subject's kind. When we apply this to Feldman and Conee's example of Jones it should clear that this rules out all sorts of possible methods of belief formation as irrelevant. Intuitively the following statements are false if we read them as classifying generics:

> Human beings forms beliefs on the basis of vision in Jones's precise environmental circumstances.
>
> Human beings forms beliefs with the content 'there is a flying saucer' on the basis of obviously defeated perceptual evidence.
>
> Human beings form beliefs on the basis of being in the exact physiological conditions that Jones is in when he forms his belief that there is a flying saucer in front of him etc. etc.

This, however, is not what allows *The Right Kind* to deal with the generality problem. Remember that the generality problem arises because for even very simple cases, the same belief can be described as the product of different methods of belief formation some of which are

intuitively reliable others of which are intuitively unreliable. As a result, reliabilist theories of justification give no clear verdict on whether a belief is justified, even for very simple cases. Now, one way of dealing with this problem is to show that, for every belief one has, there is only one method of belief formation that is epistemically relevant. If this could be done, it would solve the first aspect of the generality problem, as reliabilism would (in principle) give a clear verdict in every case. But this is not the only way of solving the generality problem. We can also solve the problem by showing that for any given belief every epistemically relevant method of belief formation will either be reliable (or not reliable). In this case, again, there is at least in principle a clear cut verdict for every case.

I adopt this second strategy in solving the generality problem for reliabilism. Here again, is *The Right Kind:*

> **The Right Kind** A method of belief-formation is relevant for the epistemic evaluation of S's belief that p, only if S is a member of kind K, and the following classifying generic is true and expresses a principled connection: Ks use M to form beliefs.[31]

Now my claim is that a method of belief formation will be epistemically relevant, only if there is a principled connection between using this method of belief formation and being a member of one's kind. Remember that when there is a principled connection between having a property and being a member of a kind, this licenses normative judgments; if the classifying generic *Ks are F* is true, and this classifying generic expresses a principled connection between being a K and being F, then an individual K will be defective it fails to be F. So a dog is defective if it has less than four legs, and a human being is (cognitively) defective if she fails to be able to form beliefs on the basis of abductive inference. Now the sorts of methods of belief formation that have a

[31] Again, with the added proviso that these sorts of methods of belief formation are truth directed.

principled connection to being a human being are such that they are all intuitively reliable: e.g., forming one's beliefs on the basis of visual perception, forming one's beliefs on the basis of memory, forming one's beliefs on the basis of various sorts of inference etc. etc.

To see this consider an unreliable method of belief formation that is widespread amongst human beings. For instance, many people are disposed to commit the conjunction fallacy: i.e., they are disposed to form beliefs that it is more probable that p and q than simply p, if they are given evidence in favor of q, but no evidence for or against p. Now this unreliable method of belief formation is wide-spread. So there is some reason for believing that the following generic is true: *human beings form beliefs about the comparative probabilities of conjunctions and their conjuncts on the basis of fallacious reasoning.* But this generic, if true, does not express a principled connection between forming beliefs about comparative probabilities in a fallacious manner, and being a human being. For the disposition to fall into the conjunction fallacy does is not connected to human nature in a way that licenses normative judgments of the following kind:

> Human beings form beliefs about the comparative probabilities of conjunctions and their conjuncts on the basis of fallacious reasoning.
>
> This human being does not form beliefs about the comparative probabilities of conjunctions and their conjuncts on the basis of fallacious reasoning.
>
> Therefore, this human being is cognitively defective.

It follows then, that if there is a true generic *human beings are disposed to commit the conjunction fallacy,* this generic expresses *a merely factual* as opposed to a *principled connection* between being a human being and committing the conjunction fallacy.

I will have more to say about this below, but this gives us reason for supposing that any true generic that expresses a principled connection between using a method of belief formation that is directed at the truth and being a human being, will be reliable.

Now, appeals to *The Right Kind* help deal with the first aspect of the generality problem, by ensuring that all relevant methods of belief formation will be reliable. But this raises some problems of its own. To see this, consider that there is a true classifying generic that connects forming beliefs on the basis of visual perception to human beings. I.e., it is true that human beings form beliefs on the basis of visual perception. Consider as well, that this method of belief formation is reliable. Going back to Feldman and Conee's example, however, Jones could be thought to form his belief that there is a flying saucer in front of him on the basis of visual perception. Now it seems we are driven to say that Jones's belief that there is a flying saucer in front of him is justified. But *ex hypothesi* it is not.

If an appeal to generics is to help with the generality problem, we need to be able to rule out this result. We can do so by focusing on the notion of an individual using a relevant, reliable belief forming mechanism in the *right way.*

What is it for an individual to use a relevant method in the right way? Again, I think appeals to classifying generics can help us answer this question. The basic idea is that for an individual to us a method of belief-formation in the right way, she must use a relevant method, in the way that her kind uses it. That is, S's forming a belief that p using a method M must be an example of a K using M to form a belief. But not just any kind of example; it must exemplify the truth that Ks use method M to form beliefs.

We can put this idea as follows:

> **The Right Way:** An individual S uses a method of belief formation M in the right way just in case: (i) S is a member of kind K, (ii) there is a true classifying generic to the effect that Ks use M to form beliefs, and (iii) S's forming a belief using M can be considered to exemplify this generic truth.

The distinction between a member of kind K being an *example* of a K using a method M to form beliefs and a member of kind K *exemplifying* the truth that Ks use a method M to form beliefs is *not* a distinction without a difference. Consider the following true classifying generic: The human has 32 teeth. Suppose that Smith has 32 teeth, but only because he originally grew 36 teeth and had four of them knocked out in a hockey fight. In this case, Smith is an example of a human with 32 teeth, but he does not exemplify the generic truth that human beings have 32 teeth. For if everyone had 32 teeth in the way that Smith did, it seems that 'human beings have 32 teeth' would no longer be a true generic statement, but rather 'human beings have 36 teeth' would be true. So an individual S's exemplifying a true generic proposition to the effect that Ks are F, goes beyond the simple fact that S is F.

I take the notion of exemplifying a generic truth as basic. But I think in most cases we can give meaningful conditions on when an individual exemplifies a generic truth. Take again a true classifying generic proposition Ks are F. If there are other true generic propositions that explain how it is that Ks are F, then an individual S exemplifies the generic truth, Ks are F, just in case S exemplifies those generic truths that explain how it is that Ks are F. So for instance, it is true that human beings have 32 teeth; but there are also truths concerning how it is that human beings come to have 32 teeth; these are the facts concerning our dental development. Smith then, has 32 teeth, but he fails to exemplify the generic truth that human beings have 32 teeth, because he fails to have 32 teeth in the way that human beings have 32 teeth; his dental development went along another path, and it is only by accident that he has 32 teeth.

When we apply this to Jones's belief that there is a flying saucer in front of him, we get the desired result: Jones bases his belief that there is a flying saucer in front of him on his visual perceptual experience, but his doing so does not *exemplify* the generic truth that human beings base beliefs on their visual perceptual experience. We can see this by noting the aspects of how Jones forms his belief that make Jones's belief lack justification, namely, that he forms his belief in the face of an obvious defeater for the belief that there is a flying saucer in front of him. It is clear, however, that when we give a complete description of how human beings forms beliefs on the basis of visual perception, part of this total description will include how beliefs based on perceptual experience are sensitive to defeaters. Jones, however, is completely insensitive to any defeaters for his perceptual experience. So while forming one's belief on the basis of visual perception is the right kind of method to be relevant for evaluating Jones's belief, Jones doesn't use it in the right way for it to be relevant.

My claim at the beginning of this section was that an individual S's belief is justified just in case S forms this belief using a relevant, reliable method in the right way. I have filled in needed detail to make this claim intelligible. A method M is relevant for evaluating an individual S's belief, only if S is a K and there is a true classifying generic that expresses a principled connection between using M and being a K. An individual S uses a method M in the right way, on the other hand, just in case her using M exemplifies the generic truth that Ks use M. Usually this will involve S exemplifying a number of classifying generics that describe how it is that Ks use M. I have also argued that this version of reliabilism avoids the generality problem without having any obviously absurd consequences. On the one hand, by making only those methods of belief formation that have a principled connection to our kind relevant, I have ensured that all relevant methods of belief formation are reliable. On the other hand, by

requiring the one forms her belief using the relevant methods of belief formation in the right way, I have ensured that my view gives the right verdict in cases in which one's belief is unjustified by formed using a relevant method. So I have given a reliabilist theory of justification that avoids the first aspect of the generality problem.

6.3.2 Habituals and the Second Aspect of the Generality Problem

In the last section, I argued for a version of reliabilism that avoids at least the first aspect of the generality problem. According to my view, a belief is justified just in case it is formed by a relevant, reliable, method of belief formation in the right way. I spelled out what it is to be a relevant method of belief formation and what it is to use such a method in the right way by appeals to classifying generics. In this section, I want to spell out what it is for a method of belief formation to be reliable in such a way that solves the second aspect of the generality problem. In particular I want to argue for the following claim:

> **Habitually Right –** A belief formation method M is reliable (in circumstances C) just in case the following habitual is true: Ks use M to form *true* beliefs.

On my view, we need to understand reliability in terms of habitual statements. Habituals are statements concerning the behavior of either an individual or a kind. Habitual statements have the following form "S Ms (when C)" or "ks M (when C)" (where C specifies some conditions) (cf. Fara (2005).) Like classifying generics, habituals admit of exceptions. Take, for instance, the habitual "John drinks wine with dinner". This habitual might be true, even if some nights John does not drink wine with dinner. It might even be true, if the majority of evenings John does not drink wine with dinner. At the same time, "John drinks wine with dinner" might be false, even if John has drunk wine with dinner on a number of occasions.

Like classifying generics, we can think of habituals as making *non-quantificational generalizations*. They make generalizations, in the sense that endorsing them involves making commitments as to what will happen through some expanse of time. But, these generalizations are not fundamentally about how many times a kind of event has occurred (or will occur), and so are not quantificational. Instead it seems that habituals, like classifying generics can be understood as often pointing out a principled connection between an individual or a kind and some behavior. As evidence of this, consider that habituals license something like formal explanations. Just as we can explain why some individual has a property by referencing the kind it belongs to, we can why an individual does something, by referencing the fact that the individual habitually does this thing. Furthermore, note that merely accidental connections between an individual and some sort of behavior are not captured by habitual statements. Perhaps, as a matter of happenstance, every time Sally smokes it has been on an odd day of the month. The habitual, "Sally smokes on the odd days of the month" however is false.

Habitually Right then amounts to the claim that the reliability of methods of belief formation is not fundamentally a matter of how many true beliefs a type of method produces. This idea might smack of paradox. How could reliability be something other than a matter of how many true beliefs a method brings about? However, the claim that the reliability of methods is not fundamentally a matter of how many true beliefs a type of method produces is well-motivated.

First, it seems that the primary notion of reliability has to do with idea of being dependable, i.e., of being worthy of being relied upon (cf. Schafer (2014) and Turri (Forthcoming). Of course, to be worthy of being relied upon to X in conditions C, one must generally or habitually X in conditions C; but there is no reason for supposing that the idea of

being dependable or worthy of being relied upon carries with it presuppositions in regards to how often or at what rate one must be successful.

Second, note that we are happy to attribute reliability to any number of belief-forming processes for which we have no clear notion of their rate of success. Are the general methods of philosophical or scientific research reliable? It seems so. But what is the ratio of true to false beliefs produced by these methods? We have no way of knowing this. How many beliefs have been formed by these methods? By how many different people? What exact beliefs have been formed by these methods? We lack the means, it seems, to answer these questions. And even if we could answer these questions, it seems that it doesn't matter. No matter what the ratio of true to false beliefs produced by these methods actually is, we will accept these methods as reliable. Again, this might strike us as paradoxical, but note that Goldman's definition of reliability seems to admit this. For Goldman, a belief-forming process is reliable if it has a high truth ratio; but a belief-forming process has a high truth ratio for Goldman just in case it meets some threshold criterion, which he allows to vary. In principle it seems it can vary so much that the truth ratio of some process could be considered to be high, but under .5 (Cf. Turri Forthcoming).

Given this, there seems to be little reason for demanding that we understand the reliability of belief forming processes in terms of the ratio of true to false beliefs that they produce. Rather, there is strong reason for understanding reliability in the way that I have suggested, i.e. in terms of habituals. On the one hand, there is no pre-theoretical reason for understanding reliability in terms of truth-ratios. And on the other hand, trying understand reliability in terms of truth ratios necessarily brings about the second half of the generality problem. Recall that the second half of the generality problem arises from needing to specify a reference class in order to determine the truth-ratio of a given belief forming mechanism. The

difficulty seems to be that there is no principled way of specifying the reference class, and so reliabilism is incapable of giving us a determinate verdict regarding even the simplest cases. If we understand reliability in terms of habituals, however, the second part of the generality problem never arises. Given that habituals are non-quantificational generalizations, there is no need to understand this habitual statement in terms of the ratio of true beliefs to total beliefs produced by the relevant method. As a result there is no need to specify any reference class to determine the truth of the claim that the method is reliable.

If I am right, then, we have no reason for supposing that reliability needs to be spelled out in quantificational terms; furthermore, we have reason for wanting to specify reliability in terms of a non-quantificational generalization, since understanding reliability in this way allows us to avoid the second half of the generality problem for reliabilism. This is exactly what appeals to habituals allow us to do.

Before moving on, it might be helpful to go back to the list of methods of belief formation I proposes to have a principled connection with the kind *human being*. To solve the first aspect of the generality problem, I made the claim that the methods of belief formation that have a principled connection with the kind *human being* were all reliable. These methods included the following: forming beliefs on the basis of perceptual experience, forming beliefs on the basis of memory, forming beliefs on the basis of inductive, deductive, and abductive inference etc. etc. We can now see that *Habitually Right* bears out my earlier claim. For it is true that human beings form true beliefs on the basis of perceptual experience, memory, inductive, deductive, and abductive inference etc. It follows then, that these methods of belief formation are reliable.

6.4 OBJECTIONS AND REPLIES

In the past sections I have laid out a version of reliabilism that successfully deals with the generality problem. According to my view, an individual S's belief is justified just in case it is produced by a relevant, reliable method of belief formation in the right way. A method of belief formation is relevant for evaluating an individual's belief only if there is a principled connection between the kind to which the individual belongs and using this method of belief formation. The individual uses a method of belief formation in the right way just in case her using the method exemplifies the generic truth that members of her kind use the relevant method of belief formation. Finally, a method of belief formation M is reliable just in case there is a true habitual 'members of S's kind use M to form true beliefs'.

I believe that my version of reliabilism solves the generality problem. But one might object that it falls prey to another well-known problem for reliabilist theories of justification, the *new evil demon* problem. We can put the problem in the following form:

1.) Non-skepticism: Many of the beliefs that we have are justified.
2.) The possibility of skeptical scenarios: It is metaphysically possible that we exist in a world in which we possess all the same non-factive mental states, but in which we are in a radical skeptical scenario in which the great majority of our beliefs are false.
3.) Parity: The beliefs of our counterparts at these skeptical worlds are justified if our beliefs are justified.
4.) Reliabilism: An individual's belief that p is justified just in case it is produced by a reliable belief forming mechanism.

The new evil demon problem, clearly poses a difficulty for simpler forms of reliabilism, that hold a belief is justified just in case it is produced by a reliable method of belief formation, where reliability is understood in statistical terms. In this case, propositions 1-4 are mutually inconsistent.

My version of reliabilism, of course, does not understand reliability in statistical terms. But it still might seem difficult for it to accommodate our intuition that our counterparts at skeptical worlds have justified beliefs. On my view, an individual S's belief is justified only if her forming her belief using method M exemplifies a true classifying generic of the following form 'Ks use M to form beliefs". Furthermore, S's using a method of belief formation exemplifies the true classifying generic only if S uses the method in the way that her kind does. Now, when we spell out how it is that members of our kind use various methods of belief formation, it seems that this description will have both internal and external components. So for example, it seems that the complete description of how human beings form beliefs on the basis of perceptual experience will include a description of how human beings are sensitive to evidence they already possess when they form beliefs on the basis of experience. But it will also include a description of the environments in which the human being forms beliefs on the basis of perceptual experience.

Now, it is clear that our counterparts at skeptical worlds are not using the relevant methods of belief formation in the right way, in so much as they are not in the right sorts of environments. Human beings form beliefs on the basis of perceptual experience, and they do so in environments quite different than the sorts of environments in which our counterparts at skeptical worlds find themselves. So even though my view doesn't understand reliability in statistical terms, it seems that it struggles to accommodate the intuition that the beliefs of our counterparts at skeptical are justified.

One response, might be to call into question our intuition that the beliefs of our counterparts of skeptical worlds are justified. Littlejohn (2009), for instance, argues that it is

wrong to believe that false beliefs are justified, since a belief can justified only if it can be defended, and ultimately, false beliefs cannot be defended.

Another interesting possibility is to make some slight revisions to the view I have been defending. As noted above, my view rules that our counterparts in skeptical scenarios fail to have justified beliefs, because when they form beliefs on the basis of e.g. perceptual experience, they do not exemplify the generic truth that human beings form beliefs on the basis of perceptual experience. The reason is not because of any internal, strictly psychological component, but rather because of environmental conditions in which our skeptical counterparts find themselves. This suggests the following amendment:

> **The Right Way *** An individual S uses a method M in the right way just in case (i) S is a K, and (ii) S's using M either exemplifies the truth that Ks use S to form beliefs, or S's using M fails to exemplify this truth only because of external/ environmental conditions.

The Right Way* gives us the desired results vis-à-vis skeptical cases. In these cases, one's belief is justified, because while the individual in the skeptical scenario fails to exemplify the relevant classifying generic, she does so only because the right kind of environmental/external conditions fail to obtain. We can contrast this with the case of Jones and his belief that there is an alien UFO in front of him. In this case, Jones fails to exemplify the relevant classifying generic, e.g. human beings form beliefs on the basis of visual experience, because Jones fails to take into account prior evidence in the way that human beings do when forming their beliefs on basis of perceptual evidence. So my view, with slight modification can solve the new evil demon problem for reliabilism, without having to deny that our counterparts' beliefs lack justification or making claims regarding the metaphysical possibility of radical skeptical scenarios.

6.5 CONCLUSION

In this paper I have presented a reliabilist theory of justification that avoids both aspects of the generality problem, and with minor modifications, the new evil demon problem. Of course, there is a way in which the theory of justification I have given is still just a skeleton of a theory. For the theory I have given relies heavily on true classifying generics concerning how human beings form beliefs. I have assumed throughout the paper the truth of a number of such generics. To give a full theory of justified belief, however, it would require that these claims be substantiated. This in turn would require something like the epistemological ethnography of the human being. To provide such an ethnography is a daunting task, but one I hope to have shown worth taking up.

6.6 WORKS CITED

Alston, W. (1995). "How to Think about Reliability" *Philosophical Topics* 23:1-29.

Beebe, J. (2004). "The Generality Problem, Statistical Relevance, and the Tri-level Hypothesis" *Nous* 38:177-195.

Bishop, M. (2010). "Why the Generality Problem is Everybody's Problem" *Philosophical Studies* 151:285-298.

Bonjour, L. (1985). *The Structure of Empirical Knowledge* (Cambridge: Harvard University Press).

Carlson, G. (1982). "Generic Terms and Generic Sentences" *Journal of Philosophical Logic* 11: 145-181.

----. (2010). “Generics and Concepts” in *Kinds, Things, and Stuff* ed. F. Pelletier (Oxford: Oxford University Press) pp. 16-35.

Comesana, J. (2006). “A Well-Founded Solution to the Generality Problem” *Philosophical Studies* 129:27-47.

Fara, M. (2005). “Dispositions and Habituals” *Nous* 39:43-82.

Feldman R. and Conee, E. (1985). “Evidentialism” *Philosophical Studies* 48:15-34.

Goldman, A. (1986). *Epistemology and Cognition* (Cambridge MA: Harvard University Press).

----. (2012). “What is Justified Belief” in *Reliabilism and Contemporary Epistemology* (Oxford: Oxford University Press) pp. 29-49.

Greco, J. (2012). “A (Different) Virtue Epistemology” *Philosophy and Phenomenological Research* 85: 1-26.

----. (2010). *Achieving Knowledge* (Cambridge: Cambridge University Press).

Leslie, Sarah-Jane. (2007) “Generics and the Structure of the Mind” *Philosophical Perspectives* 21: 375-403

----. (2008). “Generics: Cognition and Acquisition” *Philosophical Review* 117: 1-47.

Lepock, C. (2009). “How to Make the Generality Problem Work for You” *Acta Analytica* 24:275-286.

Liebesman, D. (2011). “Simple Generics” *Nous* 45: 409-442.

Littlejohn, C. (2009). “The Externalist’s Demon” *The Canadian Journal of Philosophy* 39:399-434.

Pelletier, F. ed. (2010). *Kinds, Things, and Stuff* (Oxford: Oxford University Press).

Plantinga, A. (1993). *Warrant: The Current Debate* (Oxford: Oxford University Press).

Prasada, S. (2010) “Conceptual Representation and Some Forms of Genericity” in *Kinds, Things and Stuff* ed. Pelletier, (Oxford: Oxford University Press) 36-59.

Prasada S. and Dillingham E. (2006). “Principled and Statistical Connections in Common Sense Perception” *Cognition* 99:73-112.

Schafer, K. (2014). “Knowledge and Two-Forms of Accidental Truth” *Philosophy and Phenomenological Research* 89: 373-393.

Swinburne, R. (2001). *Epistemic Justification* (Oxford: Oxford University Press).

Turri, J. (Forthcoming) “Unreliable Knowledge” *Philosophy and Phenomenological Research*.

Wunderlich, M. (2003). “Vector Reliability: A New Approach to Epistemic Justification” *Synthese* 136:237-262.