Getting It Right By Accident

Masahiro Yamada*

This is the peer reviewed version of the following article: "Getting It Right By Accident". *Philosophy and Phenomenological Research*, 83: 72–105, which has been published in final form at https://doi.org/10.1111/j.1933-1592.2010.00416.x. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

> For any sentential value of *p*, (at a time *t*) a man knows that *p* if and only if (at *t*) it is not at all accidental that the man is right about its being the case that *p*.

This is Peter Unger's proposal (1968). One might hesitate about the claim to sufficiency of this but the necessity part is certainly correct. It better not be an accident that the epistemic agent's belief is correct if the belief is to have any chance of being knowledge: if someone correctly believes that *p*, pointing out that she was lucky in getting it right is a decisive objection to attributing knowledge that *p* to the agent. Let us call this requirement that it not be an accident that one is right about its being the case that *p* the *non-accidentality condition*.

* masahiro.yamada@cgu.edu

Let me clarify a few things to avoid confusion. Sam buys a lottery ticket and he wins. Of course, he is lucky to have won but this does not prevent him from knowing that he won—he could see it on the news, read it in the newspapers or come to know it in any number of other ways. Jim might stumble upon a celebrity as he steps out of a cab. Had the cab stopped a few meters away, or a few seconds earlier or later, Jim would not have spotted the celebrity and thus he would not have known that the celebrity was at this particular location. Again, this kind of luck does not prevent anyone from knowing. In such cases, one is lucky to know that *p* but it is not a matter of luck that one correctly believes that *p*. Or, in Unger's terminology, it is an accident that one knows that *p*, but it is *not* an accident that one is right about its being the case that *p*. Depending on how the world is, if one believes that *p*, one thereby believes correctly that *p* or one thereby believes *in*correctly that *p*. One can sometimes be lucky to end up believing correctly that *p* by believing that *p*; at other times, it is not a matter of luck that one ends up believing correctly that *p* by believing that *p*. The kind of accident, or luck, that gets in the way of knowledge is this: if it is accidental that one correctly believes that *p* by believing that *p*, then one does not know that *p*.

But this does not get us very far. Just under what conditions is it an accident that one correctly believes that *p* by believing that *p*? We need an answer to this question to understand what kinds of accidents are incompatible with knowledge. Unfortunately, Unger has very little to say about this. Quick reflection can show that our assessment of when our achieving something by doing something is accidental is rather complex and not identical with the notion of a fluke. For instance, I turn on the light and thereby alert the burglar. It could be just plain luck—an accident—that I ended up alerting the burglar. But this is not quite to say that it

was some kind of fluke that it happened. Typically, the conditional probability of my alerting the burglar given my turning on the light is very high—a typical burglar will be alerted by lights turning on. The notion of luck, or accident, is more complex than our notion of probability.[1] The aim of this paper is to delineate the conditions that must be satisfied if one's belief is to *not* count as true by accident in a way that is incompatible with knowledge.

Fortunately, we can extract some ideas from various proposed necessary conditions of knowledge that are supposed to supplement or replace the traditional justified true belief account of knowledge. The various counterexamples to the justified true belief account of knowledge work by examining situations in which the epistemic agent has a justified true belief that p but intuitively does not know that p. In such situations, it is quite natural to say that even though the agent is justified in believing that p, it is still just an accident that s/he got it right. Thus, in Bertrand Russell's example of a man correctly believing that the current British Prime Minister's name begins with a 'B' because he believes that the name is Bannerman, it is clearly just an accident that the man got to believe correctly that the Prime Minister's name begins with a 'B' (Russell 1997, 131);[2] the same goes for the more complex examples given by Edmund Gettier (1963). Similarly, in Goldman's fake barn county, Henry is lucky to correctly believe that the thing he is looking at is a barn (Goldman 1976).

Theories of *knowledge* that are developed as reactions to such counterexamples

[1] This was already noted by Aristotle in his discussion of coincidence and luck in his *Physics* II.4–9.

[2] Russell's own concern is with showing that knowledge that p is more than a true belief that p so that it is not intended as a counterexample to the justified true belief account of knowledge. But the point remains that one can end up with a true belief as a matter of luck in which case one does not know.

to the 'traditional' account of knowledge can point the way in which we need to look in order to find the right account of the non-accidentality condition. I will discuss briefly three such accounts of knowledge. In discussing the benefits and shortcomings of these views, we can also find some constraints on the account of the non-accidentality condition that I am after. The views I want to discuss briefly are: process reliabilism, sensitivity and safety accounts. I will then proceed to provide my own take on the non-accidentality condition. One of the important results will be that knowledge requires neither sensitivity nor safety.

1.1. Process Reliabilism

According to process reliabilism, a belief is knowledge only if it has been formed through a reliable belief forming process (Goldman 1979). A process is reliable just in case it is likely to produce a true belief. If a belief has been formed through a reliable process and it is true, then it is usually no accident that the belief is true. This is part of what it means to say that a belief forming process is *reliable*. So perhaps we can understand the non-accidentality condition as the condition that the belief must be formed via a reliable belief forming process. But this understanding of the non-accidentality condition faces a difficulty. Consider the following belief forming process: Jim buys a lottery ticket; he knows that there are n tickets and that there is only one winning ticket; he reasons 'there are n tickets and one winning ticket; so the chance of my winning is $1/n$; so I will not win the lottery.' The larger the number of tickets n, the more reliable this piece of reasoning. However, no matter how large n is, we are disinclined to attribute to Jim the knowledge that he will not win. The reason appears to be that there is always some chance that he will win and hence there is some level of accidentality left in his correctly believ-

ing that he won't win. As Unger aptly puts it, it must be *not at all* an accident that one gets it right. On the other hand, we readily grant knowledge on the basis of memory even though we know full well how faulty and of limited reliability our memory faculties are. This indicates that we do not regard it as an accident when memory works since otherwise we would not accept a belief formed on the basis of memory as a piece of knowledge (satisfaction of the non-accidentality condition is *necessary* for knowledge). An adequate understanding of the non-accidentality condition must be able to account for this somewhat odd phenomenon.[3] Let us next look at an analysis of knowledge that satisfies this constraint.

1.2. Sensitivity

According to the sensitivity, or tracking, account proposed by Nozick (1981), S's true belief that p is knowledge if and only if a) if p were not the case, S would not believe that p; b) if p is the case, S would believe that p.[4] Put informally, this means that S's assessment of p must be *sensitive* to the truth-value of p; or S must *track* the truth-value of p. If one is sensitive to the truth, it surely is no accident if one's belief turns out to be true. So we can think of the sensitivity requirement as one way of understanding the non-accidentality condition.

[3] Unger notes this phenomenon as well (1968, 161–2).

[4] The latter condition is often put as a counterfactual: if p were the case, S would believe that p. But this is not the most felicitous understanding of Nozick's view. First of all, if p *is* true and S does believe p (i.e. S has a true belief that p), the counterfactual turns out to be trivially true given the standard semantics for counterfactuals but the condition is obviously not meant as so easily satisfiable. Secondly, Nozick himself writes "the subjunctive [if p were true, q] is true when (roughly) in *all* those possible worlds in which p holds true that are closest to the actual world, q also is true" (1981, 173; my emphasis). I hope my formulation of b) does better justice to Nozick's intent.

The crucial point to note is that to be sensitive is not to be such that one never, under any and all circumstances, incorrectly believes that *p*. What is needed is that in those closest possible worlds in which not-*p*, one does not believe that *p*. As Nozick notes, this sensitivity condition can handle the original Gettier cases rather easily (1981, 173). Here is one Gettier-type case. I have two co-workers in the office such that I have plenty of evidence that one of them, Jones, owns a Ford and no evidence that the other owns a Ford. Based on my belief that Jones owns a Ford, I deduce that someone in the office owns a Ford. It is in fact true that someone owns a Ford but it is not Jones: it is the other co-worker who in fact owns a Ford. My belief that someone in the office owns a Ford is both justified and true but it is not knowledge. The sensitivity condition can explain this nicely: I do not know that someone owns a Ford because, given the circumstances, I would believe that someone owns a Ford even if no one did—the closest possible world in which no one in the office owns a Ford is still a world in which I believe that someone owns a Ford since in that world I still have all the evidence that indicates that Jones owns a Ford.

The sensitivity condition can also handle the phenomenon that posed a problem for process reliabilism above. In the case of the lottery, the closest possible world in which Jim holds a winning ticket is still a world in which Jim would believe that he is not going to win since in that world the odds remain the same and from Jim's point of view, nothing relevant is different. So Jim does not know that he is not going to win. On the other hand, Jim knows that he had a fried egg for breakfast on the basis of memory since in the closest possible world in which he did not have a fried egg for breakfast, it would not seem to him that he had a fried egg.[5]

[5]This assumes, of course, that Jim has memory faculties that function normally. If he were under

So far so good. However, the sensitivity account faces a serious problem. We want our theories of knowledge to at least allow for the possibility that we know many ordinary facts such as that we have hands, that there are trees and the like. The trouble is that a non-skeptical epistemological theory that takes sensitivity to be a necessary condition of knowledge must abandon the closure of knowledge under known entailment. Take propositions p and q such that p entails q. Suppose Jones knows that p and knows that p entails q. The sensitivity requirement says that in order to know that q, Jones must be sensitive to the truth-value of q. But this can fail. For example, suppose Jones knows that he is sitting on a chair. His belief that he is sitting on a chair satisfies the sensitivity condition (if he were not sitting on the chair, he would not believe it; if he sits on a chair, he would believe it). He knows that if he is sitting on a chair, he is not a brain in a vat who is being fed experiences as of sitting on a chair (since a brain in a vat cannot sit; it's just a brain floating in some liquid). He, of course, also believes that he is not a BIV fed certain sorts of experiences as of sitting on a chair. But he does not know that he is not a BIV: if he were a BIV being fed experiences as of sitting on a chair, he would believe that he is *not* a BIV being fed experiences as of sitting on a chair. His belief that he is not a BIV being fed certain sorts of experiences is not sensitive. Thus, this belief is not knowledge. This is so even if Jones realizes that his not being a BIV fooled into thinking that he is sittng on a chair is entailed by the two propositions a) that he is sitting on a chair and b) that this necessitates that he is not a BIV. Jones can know both these propositions, realize that they entail that he is not a BIV and yet fail to know that he is not a BIV.

---

the constant influence of memory altering drugs, say, he would not count as knowing that he had a fried egg for breakfast even if he correctly remembers that he did.

More generally, on a non-skeptical epitemological theory that takes sensitivity to be a necessary condition on knowledge, knowledge turns out not to be closed under known entailment: S can know that p, know that p entails q, and yet be in no position to know that q. Of course, if one is willing to hold that we cannot know many of the ordinary facts we take ourselves to know, it is possible to hold that both sensitivity and closure of knowledge under known entailment are necessary conditions for knowledge (see Warfield 2004). For present purposes, I shall take it for granted that skepticism is not acceptable. But then the remaining alternative is to reject closure of knowledge under known entailment.

The denial of closure allows us to insist that we do know many of the things we commonly take ourselves to know—e.g. that there are many people, many objects in my surrounding, etc.—while also conceding to the skeptic the intuition that we do not know that radical skeptical hypotheses are false. However, it is not just radical skeptical hypotheses that we do not know if sensitivity is a necessary condition for knowledge. It is all too easy to construct cases in which closure of knowledge fails if the sensitivity condition is accepted. A general recipe for this is: let p entail q and let S be sensitive to p; construct a situation such that if q were not true, r would be true which would make S believe q anyway. For instance, suppose I am looking at a cow lying in the grass and that hidden from sight by the cow is a decoy duck that I cannot distinguish from a real duck. I believe there is a cow and I know it; since a cow is an animal, I deduce that there is an animal; this latter belief is not knowledge because if there were not an animal there, i.e. if the cow were not there, there would be the decoy duck in sight and I would still believe, falsely now, that there is an animal (though I would not believe that there is a cow; I would believe there is a duck). Hence, according to the sensitivity

account I know in this situation that there is a cow while not knowing that there is an animal. We can easily multiply such examples that do not involve any radical skeptical hypotheses. DeRose aptly characterizes this consequence as *abominable* (1995, 27–8).

The problem is not that sensitivity does not guarantee that deduction from known premises always extends knowledge. Recent discussions of so-called transmission failure provide ample documentation that deduction from known premises does not always extend knowledge (e.g. Davies 2004, Pryor 2004, Wright 2002). Transmission failure can happen when one already knows the conclusion or when one knows the premises only because there is independent support for the conclusion which in turn supports the premises. In such cases, closure does not fail. One knows (or is in a position to know) both the premises and the conclusion, or one knows (or is in a position to know) neither. But consider the example of the failure of knowing that one is not a brain in a vat. According to a non-skeptical account of knowledge that takes sensitivity to be a necessary condition of knowledge, one can know that one has hands, arms, friends, etc., know that this entails that one is not a brain in a vat, and yet there is *no way* for one to know that one is not a cleverly deceived brain in a vat. Skeptical hypotheses are constructed in such ways that one could not be sensitive to the truth/falsity of the hypotheses so that one *could not* possibly know whether or not they obtain if the sensitivity account is correct. What is abominable is that it could be that there is *no way* to know that q even though one knows that p and one knows that p entails q.

Since it is a necessary condition of knowledge that the belief not be true by accident, if we explicated the non-accidentality condition in terms of sensitivity, we would be committed to treating sensitivity as a necessary condition for knowledge.

As there is ample reason to avoid this commitment—it either leads to skepticism or the denial of closure under known entailment—we must look for other ways of understanding the accidentality at issue.

1.3. Safety

Safety is a close relative of sensitivity. Prominent defenders of safety include Sosa (1999) and Williamson (2000, esp. chs. 4 and 5). The basic idea of safety is: a true belief that p is safe just in case it would not be so easy to be mistaken about p. And if it is not so easy to be mistaken about p, it would not be an accident if one gets it right that p. More formally, the idea is that a belief that p is safe just in case in nearby possible worlds in which one believes that p, p is also true. If not-p is a relatively 'realistic' possibility, safety requires sensitivity: in nearby possible worlds in which it is not the case that p, one must not believe that p for otherwise it would be easy to be mistaken about p. However, for those p which are such that not-p worlds are far-fetched possibilities, safety does not require sensitivity.

For instance, here is an example given by Sosa (1999). If I throw a garbage bag down a garbage chute, I know that it has arrived downstairs in the garbage bin. This is so despite the fact that if something were to go wrong and the bag of garbage were to get stuck on its way down, I would still believe that the bag arrived downstairs. My belief that the bag arrived downstairs is thus not sensitive. Any view that has sensitivity as a necessary condition for knowledge will have to say that in this case I do not know that the bag arrived downstairs. A sensitivity theorist might try to blunt the force of this kind of counterexample by holding that one does not in fact know that the bag arrived downstairs. This is the position held by Becker (2007). At most one knows, according to Becker, that it is very likely

that the bag arrived downstairs. He suggests that this is supported by 'logic textbooks, almost all of which define an inductive argument as one that makes the conclusion *probably* true.'(2007, 55). But this cannot be the right interpretation of the point made by logic textbooks. An inductive argument for the conclusion that *p* is not an argument for the *different* conclusion that *it is probably true that p*. If logic textbooks are making any claims about *what* one knows, it has to be the claim that if an inductive argument can generate knowledge for the conclusion, it does so despite the fact that the conclusion is not entailed by the premises—i.e., despite the fact that it is only probably true given the premises. So, if anything, the logic textbooks support, *pace* Becker, the commonsense view that one *can* know that the bag arrived downstairs, and not merely that it is very likely that the bag arrived downstairs. The safety view accounts for this by noting that given that the chute is in proper working order it is not so easy to be mistaken about whether the bag arrived downstairs: the closest possible world in which the bag fails to arrive downstairs is rather far away from the actual world.[6]

Similarly, it is not so easy to be mistaken about one's hands being real rather than being superbly engineered prosthetics that are indistinguishable from the real thing to the naked eye. After all, in order to be mistaken about one's hands being real rather than superbly engineered prosthetics, one would have to be in a rather far-fetched situation. Similarly, it is not so easy to be mistaken about one's not

[6]Becker notes that it would be possible to know that the bag arrived downstairs if one would notice it were the bag to fail to arrive downstairs because, say, the chute would make a funny noise (2007, 56). This is, of course, true but does not address Sosa's concern. Clearly, the case Sosa has in mind is one in which one would not notice anything if the bag were to fail to arrive downstairs. The question is what to say in such a case and Becker's assessment is that one does not know that the bag arrived downstairs.

being a brain in vat. Thus, according to this view, it is possible to have a safe belief that one has hands, have a safe belief that this entails that one is not a handless brain-in-a-vat, and to have a safe belief that one is not a handless brain-in-a-vat. Abominable conjunctions can be kept at bay.

Recently, Pritchard (2005) has proposed safety as an account of what he calls veritic luck. A belief is veritically lucky if "it is a matter of luck that the agent's belief is true" (2005, 146) where the belief's being true by luck is incompatible with its being knowledge. So on his view, the non-accidentality condition is simply safety.[7] But safety cannot be the non-accidentality condition as safety is not a necessary condition for knowledge. To see this, consider the following case. Garrett is a disgruntled employee of a large bank. He is out of money and even though he occasionally buys lottery tickets, he is astute enough a statistician to know that expecting to win the lottery is unreasonable. In fact, he regularly discards lot-

[7]Pritchard is not fully explicit whether her thinks safety is sufficient as an account of the non-accidentality condition. He notes that what he calls 'reflective epistemic luck' is also epistemically problematic but he is not clear whether it is really incompatible with knowledge. According to him, a belief suffers from reflective epistemic luck if "given only what the agent is able to know by reflection alone, it is a matter of luck that her belief is true" (2005, 175). To avoid circularity, being reflectively epistemically lucky must not be understood as one way in which it could be "a matter of luck that [a] belief is true." If the absence of reflective epistemic luck is a requirement for knowledge, it is best understood as a further condition on knowledge rather than a part of the non-accidentality condition. As stated, it seems to me that the absence of reflective epistemic luck is *not* a condition on knowledge since the condition amounts to the requirement that it be purely *a priori* that a given belief is true in a non-accidental way. This will lead to skepticism—as Pritchard argues, the root of philosophical skepticism is the demand for elimination of reflective epistemic luck. So it seems that Pritchard thinks that safety is sufficient for an account of the kind of accidentality that gets in the way of knowledge.

tery tickets without bothering to check whether he won. So instead of relying on the lottery system, he decides to steal money from his bank. Luckily, he has the privileges and skills necessary to carry out this theft. His plan is to plunder small amounts of money from each and every client account but to make changes to the computer system so his acts will go undiscovered long enough for him to escape to a safe place. In particular, none of the clients will see anything amiss when they check their account information online or otherwise. Today is the date for which he has been planning. On his way to work, he stops by at a gas station and decides on a whim to check his lottery ticket that has been in his pocket for a while. He has hit the jackpot. Garrett scraps his plan and does nothing to harm the bank or its clients. A few hours later, Lucy, a client of the bank, checks her bank account online and the system tells her that she has enough cash in the bank to pay her rent.

Lucy *knows* she has enough money to pay her rent. After all, nothing is wrong with the bank's computer system. But her belief that she has enough money is not safe. Given the way she forms her belief—relying on what the online banking system tells her—she could easily have formed a false belief. Garrett's winning the lottery is an extremely unlikely event and given his attitude towards lotteries, he could easily have failed to check his ticket. But if he had not found out that he won the lottery, he would have stolen money and manipulated the computer system in such a way that Lucy would have formed false beliefs about the money sitting in her account. So there are nearby possible worlds in which Lucy falsely believes that she has enough money to pay her rent; i.e. in these circumstances, it would be easy for her to form the false belief that she has enough money.[8] So Lucy's belief

[8]Neta and Rohrbaugh (2004) and Comesaña (2005) give similar counterexamples to safety.

that she has enough money to pay her rent is not safe. But it is knowledge. Thus, safety is not a necessary condition for knowledge and hence safety cannot be the non-accidentality condition that we are after.

2. The Three Aspects of Non-Accidentality

What we have seen so far indicates that the non-accidentality condition is more complex than meets the eye. The possible views discussed above all treat the non-accidentality condition as a rather simple condition. But a closer inspection of a few cases shows that the phenomenon is somewhat more complex than each of the views discussed would suggest.

Suppose Pericles consults the Oracle of Delphi to find out whether the Persian army is approaching from the north or the south. Pericles's belief might be justified if he has the right kinds of background beliefs. He could also come to the correct conclusion about the strategy of the Persians. But even if he did end up with a justified true belief about the strategy of the Persians, he would not *know* the route of the Persian army. Why? Because believing what the Oracle tells one is not truth-conducive in the sense of being at least likely to lead to one's holding a true belief. Pericles might, of course, have a justified belief that believing the Oracle is truth-conducive. But justified belief that the Oracle is truth-conducive is not enough. At the very least, believing the Oracle must *be* truth-conducive. We can conclude that the method used to answer the question whether p must be truth-conducive if one's getting the right answer is not to be an accident. But this is not enough.

A further source of accidentality is the possibility of misapplying a method. Recall the days when Saddam Hussein was still at large and it was known that he

had many doubles. Consider the following method as used in those days: if you see Saddam Hussein give a speech, conclude that he is alive! This method is quite truth-conducive: it cannot be that you see Saddam Hussein give a speech unless he is also alive.[9] However, given that he has many doubles, it is too easy to misapply the rule by seeing a double giving a speech. So the situation must be such that the method used can be reliably applied; i.e. it must not be an accident that one managed to do as the method requires.

Finally, it matters why one uses a certain method. Suppose Luke is at the gym and wants to weigh himself. There are two scales, one red and one blue, and they give conflicting reports of his weight. Luke decides to toss a coin in order to decide which scale to rely on. The toss results in his relying on the blue scale which is in fact the reliable one. From now on he always relies on this scale. His method of figuring out his weight (see what the blue scale reports) is truth-conducive, it is easy to use (i.e. little chance of misapplying it) but it is still an accident that he gets his weight right. The reason is that he started using the method in a way that has nothing to do with which scale is in fact reliable and so Luke does not know his weight.[10] It is an accident that he is using a truth-conducive method and this violates the non-accidentality condition.

So we have three aspects of the non-accidentality condition. Given a true belief that p formed via a method M, for it to be no accident that the belief is true, it must

[9] I am using the locution 'see X' in its success sense so that one cannot see X unless X is there to be seen.

[10] As Foley (1984; 1987, 204–8) has argued, it might be that there is a sense in which Luke can be rational in believing that his weight is such-and-such in the described situation depending on what else he believes. However, the present focus is on knowledge. Even if Luke is rational, it does not follow that he knows.

be the case that:

1. the method M used is truth-conducive
2. it is not an accident that one correctly applied M
3. it is not an accident that one is using a truth-conducive method

When we say that it is *not at all* an accident that S correctly believes that p, what we mean is that there is no sense in which it is an accident that S correctly believes that p. Satisfaction of one the above conditions may make it in some sense not an accident that one got it right, but it will leave other senses in which it is still an accident. For it to be not at all an accident that S correctly believes that p, all three conditions must be satisfied. The rest of this paper is devoted to clarifying what these three conditions amount to.

The talk of use of methods may strike some as inappropriate in this context as it might be taken to involve the problematic commitment that agents must have intentions to conform to certain methods in forming beliefs. Such a commitment would indeed be problematic and the talk of methods is not meant to involve it. For the purposes of this paper, an agent's using a method consists in the agent's having a complex set of dispositions such that under certain idealized circumstances he would conform to the method. Under non-ideal circumstances, an agent can use a method without conforming to it and, as it will become clear, it is this possibility of failure to conform to a method that is important for present purposes. I will have more to say about the dispositions involved in using a method later on. An important point to keep in mind is that two agents displaying the same behavior in the actual world can be using different methods. For instance, two members of the jury might come to believe the defendant guilty upon hearing a witness

assert that she saw the defendant commit the crime. But one jury member might be using the method to trust any and all witnesses, the other might be using the method to trust only witnesses of a certain kind to which the witness in question happens to belong. The difference between the two will be in their dispositions—what they are disposed to do when confronted with various kinds of witnesses—and the difference need not manifest itself in the actual world. This means that, for instance, in the case of Saddam's speech above there is a real difference between using the method I specified above and using the non-truth-conducive method 'if you seem to see Saddam give a speech, conclude that he is alive!' even if both might lead to the same behavior in the actual world. Thus, there is a difference between a belief's being true by accident because a non-truth-conducive method was used, and a belief's being true by accident because a truth-conducive method was used but was too easy to misapply. This is why we need to distinguish between the first and second aspects of the non-accidentality condition. For purposes of giving cases, I will often simply stipulate which method a given agent is using just as I will be stipulating other aspects of a case. What matters for my purposes is when a belief is true by accident *given* the agent is using a particular method.

With the above caveat about how to understand my talk of methods here, the kind of complexity in the notion of accidentality is what one would expect. Consider a plan to rob a bank. The plan can be evaluated according to two different sorts of criteria. First, how likely is success *if* one does exactly as the plan specifies? A plan might be lousy because success is unlikely even if one executes the plan exactly as specified. If one succeeds nevertheless, it would be a matter of luck. Secondly, how difficult is it to do exactly as the plan requires? Some plans are very difficult to carry out. Perhaps, the plan requires to dodge machine gun fire at close

range while carrying a bag weighing in excess of three hundred pounds. It could be that *if* one manages to do this, success is very likely. Even so, given the difficulties involved in carrying out the plan, it would be a matter of luck if one succeeds. Thirdly, one might accidentally do what the plan requires to do. For instance, one might through sheer accident, slapstick style, go through exactly the complex set of motions as specifed by the plan in order to avoid detection by the security system. If this happens, success would be a matter of luck. Given these considerations in the more general action case, the presence of the three factors outlined above in the epistemic case is exactly what we should expect.

A word of warning is in order. I do not regard the following as a *conceptual analysis* if analysis means somehow providing an explication of meaning. The following is like an attempt of a security expert trying to find out what needs to be done to avoid inadvertently letting in dangerous material inside a passenger plane. Just as the security expert is not in the business of clarifying the meaning of the locution 'inadvertently letting in dangerous material,' I am not in the business of trying to find out the meaning of 'it is no accident that one correctly believes that *p*.' Rather, I am trying to find out what conditions we need to satisfy if we are to be such that it is no accident that we correctly believe that *p*.

2.1. The Method Used Must be Truth Conducive

The canonical expression of a method is: if *C*, do *A*! In an epistemological context, the methods in question are methods for forming judgments so their canonical expression is: if *C*, judge that so-and-so! Sometimes, the phrase 'so-and-so' stands for a phrase expressing a proposition but often it does not. Consider: if you hear droplets hitting the roof, judge that it is raining! If you apply this method while in

Denver, you will judge that it is raining in Denver. Had you been in Katmandu, you would have believed that it is raining in Katmandu. So which proposition a subject ends up believing by applying this method depends on which possible world the subject is in. But not only does it depend on which possible world the subject is in, but also where the subject is within a given possible world. For instance, a subject who is in Denver will judge, using the above method, that it is raining in Denver. A different subject in the same possible world but located in Katmandu will judge that it is raining in Katmandu. I will write 'if C, judge that *P(w_s)*!' as the schematic representation of a method of judgment. w_s is a so-called centered possible world centered around the subject S and *P(w_s)* yields the proposition to be believed by S in world w dependent on S's location within w. In some cases, *P(w_s)* is a constant function yielding the same proposition for all possible worlds and subjects as in the case of the method 'if Socrates is a man and all men are mortal, judge that Socrates is mortal!' But in many cases, *P(w_s)* will yield a different proposition with different values for w and S as in the case of the method 'if you hear droplets hitting the roof, judge that it is raining!' The proposition S actually believes by using the method 'if C, judge that *P(w_s)*!' is the proposition *P(actual world$_s$)*.

Now, to say that an epistemic method is truth-conducive is to say that *if* one does as the method says, one will end up with a true belief. More precisely, given the method 'if *C*, judge that *P(w_s)*!', it must be true that in a nearby world w in which *C*, *P(w_s)* is true as well for otherwise there would be nearby worlds in which one does as the rule requires and ends up with an incorrect belief that *P(w_s)*— it must not be so easy to go wrong while correctly using the method. The flip-side of this is that if *P(w_s)* were not true, *C* would not be true in w either since otherwise one would believe *P(w_s)* even if not-*P(w_s)* by correctly judging as the

method requires—the method says to believe *P(w_s)* if *C* and it should not be so easy to go wrong about the truth-value of *P(w_s)* when one does as the method requires. Let us express these conditions by:

i. If *C*, *P(w_s)*

ii. If not-*P(w_s)*, not-*C*

If these conditions are met, then a method is truth-conducive. These conditionals are not to be understood as material implications or as entailments. They are the ordinary English conditionals as they appear in the following kind of conversation:

Marc: How do you know whether Debby is in her office or not?

Matthew: I just check whether the lights are on. *If the lights are on, she's in.*

Clearly, what Mathew means is not the material implication: he cannot prove that what he asserts is true by pointing out that the lights are off or by noting that Debby is in her office. Nor does he mean entailment: of course, there are exceptional situations—she might forget to turn off the light at the end of her day—but usually, if the lights are on, she is in. I trust the reader has an intuitive grip on the kind of conditional I have in mind.[11] One important feature of this type of conditional is that, like with many other conditionals that are not material implications

[11]For those more technically minded: 'if A, B' as I am using it here is true just in case in nearby centered possible worlds in which A, B is also true. It is not just the closest centered possible world but a whole range of close centered possible worlds that play a role in evaluating the conditional. One thing that seems clear is that the worlds must be understood as centered possible worlds: as an anonymous referee has pointed out to me, in the example given above in the text, what is the case at nearby *times* is also important in deciding whether it is true that if the lights are on, Debbie is in her office. Who or what is to be at the center of the relevant centered worlds is determined by context. Of course, the grip we have on the 'range of close

or their necessitations, the following inference pattern is *not* valid: if A, B; if B, C; therefore if A, C. That this is so will be shown shortly while discussing Dretske's famous painted mule. This is important for if the inference were valid, my final account would entail both the sensitivity and safety conditionals and hence have the disadvantages of both. However, since the inference is not valid, my account is crucially different from either of them and is capable of avoiding the counterexamples to them.

The conditions *i.* and *ii.* are neither the sensitivity nor the safety condition. The sensitivity and safety conditions concern the connection between an agent's belief that *p* and the truth-value of *p*. The conditions *i.* and *ii.*, on the other hand, concern the connection between the truths of C and *P(w_s)*. By themselves, they tell us nothing about any connections between the agent's beliefs and the truth-value of *P(w_s)*. In particular, even if there are many nearby worlds in which the agent falsely believes that *P(w_s)* by using the method, the method can still be truth-conducive. This would happen, for instance, if the rule is very easy to misapply so that in

centered possible worlds' is no better than the grip we have on whether or not a claim like 'if the lights are on, Debby is in her office' is true. This is the main reason why I will rely on our intuitive grip and will use formulations kept as closely as possible to plain English.The same referee has suggested to me that the modal import of the ordinary language conditionals can be explained away by appeal to implicatures; i.e the only semantically relevant centered possible worlds in evaluating the conditionals are centered versions of the actual world but never centered versions of non-actual worlds. This strikes me as rather unlikely though I admittedly have no knockout argument that such a view cannot be right. If it should turn out to be right, I would retreat to the position that the conditionals I am employing are to be understood in a technical sense such that their meaning is stipulated to include whatever modal import is implicated by the usage of the ordinary conditional in ordinary conversational contexts such as displayed above.

nearby worlds in which not-C and not-*P(w_s)*, the agent still believes that C and forms the belief that *P(w_s)* on that basis. I will have more to say on the possibility of misapplications of the method in the next section.

The above two conditions, *i.* and *ii.*, can explain why we cannot know that we lost in a lottery based on the odds no matter how long the odds of winning. Jim uses the method 'if the odds of your winning the lottery are long, judge that you are not going to win!' and forms the belief that he is not going to win the lottery. This method arguably satisfies *i.* but it does not satisfy *ii.* The satisfaction of *ii.* would require that if Jim were going to win, the odds of his winning would be short and this is not true.[12] So the method Jim uses is not truth-conducive in the right way. This is why we cannot know that we lost in a lottery based on the odds.

It can also show why relying on faulty indicators of truth does not lead to knowledge. For instance, if we rely in testimony and the witness is prone to making up stories, even if the witness happens to say something that is true, relying on the testimony will not result in knowledge. This can be explained by the fact that the method 'if witness says that *p*, judge that *p*!' is not truth-conducive in such a case. Similarly with using the constellation of planets to forecasts this year's crop: the method 'if the planets are in such-and-such a constellation, judge that the crop is going to be so-and-so!' is not truth-conducive.

Often, methods are combined into more complex methods. Suppose John wants to know whether a certain liquid tastes sour but does not want to actually taste it. What could he do? He could use the following method: if the liquid is acidic, judge that it tastes sour! Of course, he would need a way of telling whether the liquid

[12] This shows that, as should be expected, the conditionals I am using do not contrapose. Hence the need for both *i.* and *ii.* as conditions.

is acidic. Here we have the litmus test: if you expose a piece of litmus paper to a liquid and the paper turns red, judge that the liquid is acidic! John would first use the litmus test to see if the liquid is acidic. If it is, he will go on to conclude that the liquid tastes sour. Both methods satisfy *i.* and *ii.* By combining the two methods, John can be said to be using the method: if you expose a piece of litmus paper to a liquid and the paper turns red, judge that the liquid tastes sour! This method also satisfies *i.* and *ii.*

Notice, however, that a combination of truth-conducive methods does not always result in a new method that is also truth-conducive. Consider the following famous example. Sam is at a zoo, goes to the zebra pen, judges that the animal in the pen is a zebra based on its looks. He then deduces that the animal in the pen is not a cleverly painted mule (Dretske 1970).

Sam is using two methods; one for deciding whether the animal is a zebra, and another for deciding whether the animal is a painted mule. The first method is

> **Z1**: if the animal has such-and-such visible features, judge that it is a zebra!

the second is

> **Z2**: if the animal is a zebra, judge that it is not a cleverly painted mule!

Both Z1 and Z2 are truth-conducive. For Z1, we have *i.* if the animal has such-and-such visible features, it is a zebra and *ii.* if the animal is not a zebra, it does not have such-and-such visible features. Z2 is truth-conducive since being a zebra and being a mule are incompatible properties. However, their combination

> **Z3**: if the animal has such-and-such visible features, judge that it is not a cleverly painted mule!

is not truth-conducive. If the animal were a cleverly painted mule made to look like a zebra, then the animal would have the relevant visible features. So condition *ii.* of truth-conduciveness fails.

This case gives us the counterexample to the validity to the following inference that I mentioned earlier: if A, B; if B, C; therefore, if A, C. Dretske's example is set up such that: if (A) the animal is a cleverly painted mule, (B) it is not a zebra. The case is also such that: if (B) it is not a zebra, (C) it does not have such-and-such visible features. Yet, it is not the case that: if (A) the animal is a cleverly painted mule, (C) it does not have such-and-such visible features. Thus, as I asserted earlier, the inference pattern "if A, B; if B, C; therefore, if A, C" is not valid.

Z3's failure to be truth-conducive explains why we feel that if Sam were to conclude that the animal is not a cleverly painted mule in the described way, he would not know that it is not a painted mule even though it seems that the way he gets to believe that it is a zebra is impeccable and deduction is certainly a truth-conducive method: the conclusion is still a result of a non-truth-conducive method.

Does this not amount to the denial of closure of knowledge under known entailment? No. What the above shows is how *transmission* of knowledge fails. It does not show that it would be possible for Sam to know that the animal is a zebra via Z1 without at least having the wherewithal for knowing that it is not a painted mule. The truth-conduciveness of Z1 is, after all, merely a necessary condition for Sam's knowing that the animal is a zebra. For all that has been said so far, further necessary conditions of knowledge will preclude his knowing that the animal is a zebra without also knowing (or at least having the wherewithal for knowing) that it is not a cleverly painted mule.

The condition that the method one uses be truth-conducive—this is the condition

that *i.* and *ii.* explicate—explains quite a bit of epistemologically interesting phenomena. However, it cannot explain everything. One problem is necessary truths. Consider the method 'if a coin toss comes up heads, judge that Fermat's Last Theorem (FLT) is true!' This method is truth-conducive since FLT is a necessary truth: for any *C*, if *C*, FLT is true; and if FLT is false, not-*C*.[13] But someone who comes to believe that FLT is true by using this method does not know that FLT is true. There is still a sense in which his belief is true by accident.

Here is another problem. Recall an example I gave above. Suppose someone wonders if Saddam Hussein is still alive when he was still at large. He uses the method

SPEECH: if you see Saddam Hussein give a speech, conclude that he is alive!

This method is quite truth-conducive: it cannot be that Saddam Hussein is giving a speech unless he is also alive. However, given that he has many doubles, it is too easy to misapply the rule by seeing a double giving a speech. And we would thus not consider it knowledge even if one came to the correct conclusion by in fact seeing him give a speech. This leads us to the second element of the non-accidentality condition.

2.2. It Must Not be an Accident That One Correctly Applied M

What the above example shows is that it must not be too easy to misapply the method. What does this amount to? Let *M* be the method 'if *C*, judge that *P(*w_s*)*!' To misapply such a method is to form the belief that *P(*w_s*)* via *M* even though not-*C*. It must not be too easy for such misapplication to take place. This amounts to:

[13] This comes out trivially true because the antecedent is necessarily false.

it must be the case that in a nearby possible world w, including the actual world, in which one believes that *P(w_s)* via *M*, *C* is true. Moreover, it must be the case that in a nearby possible world w in which not-*C*, one does not believe *P(w_s)* via *M*. Let us put these as:

iii. If S believes that *P(w_s)* via *M*, *C*

iv. If not-*C*, S does not believe that *P(w_s)* via *M*

As with *i.* and *ii.*, these conditionals are neither the material implication nor entailment. These conditions can be met even when the sensitivity and safety conditions are not met. The sensitivity and safety conditions concern the connection between an agent's belief that *P(w_s)* and the truth-value of *P(w_s)*. Conditions *iii.* and *iv.*, on the other hand, concern the connection between C—rather than P(w_s)—and the agent's belief that *P(w_s)*. So if there is not the right kind of connection between C and *P(w_s)*, conditions *iii.* and *iv.* can be met even if the belief that *P(w_s)* is neither safe nor sensitive.

Now, in the Saddam Hussein example, both *iii.* and *iv.* fail. It is not true that if one believes that Saddam is alive via the method SPEECH, then one is seeing Saddam give a speech: it could just as well be a double. And even if one is not in fact seeing Saddam give a speech, it does not mean that one would not believe that Saddam is alive via SPEECH since one could be seeing a double give a speech.

Or consider the following example. Tom can visually distinguish German shepherds pretty well from poodles and other small domesticated dogs. On the other hand he cannot visually distinguish German shepherds from wolves very well. Suppose he uses the following method:

DOG: if you see a German shepherd, judge that there is a dog!

This method satisfies both *i.* and *ii.* and if Tom were to use this method in New York City, *iii.* and *iv.* would be satisfied as well (let us suppose that there are no huskies and the like in NYC). When he believes that there is a dog using this method, he is seeing a German shepherd. If he does not see a German shepherd, he would not believe there is a dog based on DOG. But suppose Tom is in an area where there are lots of wolves and very few German shepherds. He happens to see a German shepherd and concludes that there is a dog. Tom does not know, in this situation, that there is a dog.[14] The reason is that even though DOG is truth-conducive, in this situation it is too easy for Tom to misapply the method. Conditions *iii.* and *iv.* both fail. 'If he believes that there is a dog via DOG, he is seeing a German shepherd' is false since it could easily be a wolf. Similarly, even if he were not seeing a German shepherd, he might still easily believe that there is a dog via DOG since he might easily be seeing a wolf instead: it is not true that if he is not seeing a German shepherd, he does not believe that there is a dog.

For another example, consider the use of memory. Sam uses his memory to figure out what he had for breakfast. Let us say that the method used is 'if you have an apparent recollection as of having fried eggs for breakfast, judge that you had fried eggs for breakfast!' This method is truth-conducive since if he had not had fried eggs for breakfast, he would not have any such apparent recollection, and if he has such a recollection, then he did have fried eggs. It is also not easy to misapply—one typically knows when one has an apparent recollection as of having had fried eggs for breakfast. This explains why each time we use memory and get the right answer, we know the answer despite the fact that our chance of getting the wrong answer is much higher than the chance of incorrectly believing

[14]Example adapted from (Goldman 1976).

that one has not won in a typical lottery.

Here are the aspects of the non-accidentality conditions. S's belief that *p* formed by the use of the method 'if C, judge that *P(w_s)*!' where *p* is *P(actual world$_s$)* is true by accident if:

i. if *C*, *P(w_s)*

ii. if not-*P(w_s)*, not-*C*

iii. if S believes that *P(w_s)* via *M*, *C*

iv. if not-*C*, S does not believe that *P(w_s)* via *M*

It might be tempting to think that *ii.* and *iv.* entail the sensitivity condition via the following inference: if not-*P(w_s)*, not-*C*; if not-*C*, S does not believe that *P(w_s)* via *M*; thus, if not-*P(w_s)*, S does not believe that *P(w_s)*. However, as shown by the earlier discussion of the painted mule case, this inference is *not* valid. The conditions *ii.* and *iv.* do not entail the sensitivity condition.

Similarly, it might be tempting to think that *i.* and *iii.* entail the safety condition via the following inference: if S believes that *P(w_s)* via *M*, *C*; if *C*, *P(w_s)*; thus, if S believes that *P(w_s)* via *M*, *P(w_s)*. Again, this inference is not valid. The conditions *i.* and *iii.* do not entail the safety condition.

This means that counterexamples to either sensitivity or safety accounts need not be counterexamples to my account. In fact, my account can handle the counterexamples to either of these accounts. Consider again Sosa's example for showing that safety is a better candidate as a necessary condition for knowledge than sensitivity is. If I throw a garbage bag down the garbage chute, I know that it has arrived downstairs in the garbage bin. This is so despite the fact that if something

were to go wrong and the bag of garbage were to get stuck on its way down, I would still believe that the bag arrived downstairs. My belief that the bag arrived downstairs is thus not sensitive. However, it is safe since in most nearby worlds in which I believe that the bag has safely arrived downstairs this belief is true. My conditions so far also say that in this case nothing is amiss. The belief forming method used is:

> if you have thrown a garbage bag down the garbage chute, judge that it has arrived downstairs in the garbage bin!

This method is truth-conducive: if I throw a garbage bag down the chute, it arrives downstairs; and if a bag does not arrive downstairs, it has not been thrown down the chute. The method is also not easy to misapply: if I believe via this belief forming method that the bag arrived downstairs, I have thrown the bag down the chute; and if I have not thrown the bag down the chute, I do not believe via this method that the bag arrived downstairs. So conditions *i.* through *iv.* are satisfied in this case. So, unlike accounts such as Becker's (2007) that require sensitivity for knowledge, these conditions allow that one can know that the garbage bag arrived downstairs and not merely that it is very likely that the bag arrived downstairs.

Recall the counterexample to safety given earlier. Garrett, a disgruntled bank employee, is about to steal money from clients but in such a way that he will cover his tracks long enough to make his escape. In particular, checking the account balances online or at an ATM will not show anything amiss for a while. On the day he plans to carry out his theft, he checks his lottery ticket and it turns out he hit the jackpot and he scraps his criminal plans. Lucy, a client of the bank, checks her bank account a few hours later and forms the belief that she has enough money to pay her rent. In this case, Lucy knows but her belief is not safe. If this is right,

conditions *i.* through *iv.* have to be met and indeed they are. Lucy is using the method:

> if the banking system tells you that you have *n* dollars in your account, believe that you have *n* dollars in your account!

This method is truth-conducive. Since Garrett did not interfere with the system, the computers are working fine. If the system tells her she has *n* dollars in her account, she has *n* dollars in her account. If she does not have *n* dollars in her account, the system won't tell that she has. Moreoever, the method is easy to apply correctly. If she forms the belief that she has *n* dollars in her account via this method, the system is telling her that she has *n* dollars in her account. If the system does not tell her she has *n* dollars in her account, she won't believe she has via this method. So conditions *i.* through *iv.* are met.

The method Lucy uses is in fact truth-conducive even though the method Lucy uses *almost* would have failed to be truth-conducive. This is why conditions *i.* and *ii.* are met. Lucy almost would have failed to know how much money she has in her account since the method almost would have failed to be truth-conducive. But this does not get in the way of Lucy's actually knowing how much money she has in her account. If the banking system says she has *n* dollars in her account, she has *n* dollars in her account. After all, the system is working and to denying the truth of this conditional would be like demanding the system be repaired after being almost struck by lightning. If the system was almost struck by lightning but in fact was not, then it is working and if the system says that one has *n* dollars in one's account, one indeed has *n* dollars in the account. Lucy's case is no different: the banking system almost failed but in fact did not.

The difficulty for safety theory here is that intuitively it is still right to say that Lucy almost would have gotten it wrong so that her belief is unsafe. A safety theorist could, of course, appeal to the considerations just outlined and argue that Lucy's belief is safe. Such a version of the safety theory would evaluate a belief's safety along at least two dimensions: whether the method used is truth-conducive and whether the method was not too easy to misapply—it cannot just evaluate along only one of these dimensions for the same reasons that I had to admit these two dimensions. This is, of course, identical to my view save for the difference that I say that it is not an accident that the belief is true where this 'safety' theorist says that it is safe. The crucial point is that non-accidentality needs to be understood as having multiple dimensions rather than just one as suggested by the standard formulations of safety and sensitivity theories. My own view is that there are three dimensions. I have discussed two of them so far and the third will be discussed in the next section.

What about Goldman's fake barn county? I have postponed discussion of the case to use it to call attention to some of the worries one might have with the reliance on belief forming methods in my account of the non-accidentality condition. Goldman's case is simple enough. Henry is driving through the country side and upon seeing an object that looks like a barn forms the belief that there is a barn. It indeed is a barn but unbeknownst to Henry there are many mere barn facades in the vicinity which are indistinguishable from real barns by casual visual inspection from inside the car. Given this circumstance, Henry does not know there is a barn even though he is seeing a real barn and not a barn facade because, intuitively, it is just an accident that Henry got it right.

If the aspects of the non-accidentality condition provided so far can explain why

Henry does not know, there has to be a method that Henry is using in this case which is either not truth-conducive or easy to misapply (or both). There are at least two natural candidates:

> **HENRY-1:** If you see a barn, judge that there is a barn!
>
> **HENRY-2:** If you see something with such and such visible features, judge that there is a barn!

HENRY-1 is truth-conducive since you cannot see a barn unless there is a barn: if you see a barn, there is a barn and if there is no barn, you do not see a barn. However, HENRY-1 is easy to misapply in fake barn county: it is too easy to judge that there is a barn using this method even if one is not in fact seeing a barn. Conditions *iii.* and *iv.* fail so that if Henry is using HENRY-1, the reason he does not know is that the method he uses is too easy to misapply. HENRY-2 is not truth-conducive in the circumstances of fake-barn county: it is not the case that if you see something with barn-like visible features, there is a barn. Thus, if Henry is using HENRY-2, his problem is that he is using a non-truth-conducive method. Either way, Henry's belief that there is a barn turns out to be true by accident.

But which of these two methods is the one that Henry is using? Without some way of answering this question it is unclear whether all of the conditions *i.* through *iv.* are needed. Could one not redescribe the cases discussed so far in such a way that in each the method is not truth-conducive? Or in such a way that in each case the method is too easy to misapply? Is there a way of distinguishing an agent's using one method from her using another?

Elsewhere I argue in detail for what I take to be necessary conditions for following a rule (2010). Let me here simply state the results.

S follows the rule 'if C, do A!' only if:

1) there is a state s such that

a) S's tokening of s disposes S to do A

b) a) is true because when the disposition came into place it was the case that if S is in s, C obtains.

and

2) if S tokens the belief that C, this belief disposes S to do A.

Given these conditions, the state s is such that it disposes the agent to do A and when the disposition was set up in the agent s was also correlated with C so that the state s allowed S to conform to the rule 'if C, do A!' under the conditions prevailing when the disposition was set up. If the current conditions are different from those initially prevailing conditions, the agent may fail to conform to the rule. Condition 2) is only a conditional so that an agent need not form the belief that C obtains for her to count as following the rule as s need not be a belief. But when she believes the condition is met, she will be disposed to do A. Forming false beliefs about whether the condition obtains will be another source of failure to conform to the rule. Given these conditions, whether an agent is following a particular rule is largely an empirical matter since it depends on the causes behind the agent's relevant dispositions.

We can apply these conditions to belief forming methods. To use a belief forming method 'if C, believe that *P(w_s)*' is to follow this rule. So consider the rules HENRY-1 and HENRY-2. Assuming that Henry is a fairly normal human being, we can assume that he is such that if he believes that he is seeing a barn, this disposes him to believe that there is a barn. He is also such that if he believes that he is seeing

something with such-and-such visible features, this disposes him to believe that there is a barn. So condition 2) for rule-following is met by both these methods. What about condition 1)? Normally, Henry will form the belief that there is a barn simply as a response to his perceptual state as of seeing a barn. So there is a state, viz. the perceptual state as of seeing a barn, which disposes him to believe that there is a barn. Why does the perceptual state dispose him to believe that there is a barn? I submit that the reason a perceptual state as of *p* disposes one to believe that *p* because perceptual states are normally reliable indicators of facts around us. This means that the perceptual state as of seeing a barn disposes Henry to believe that there is a barn because the state as of seeing a barn is correlated with his actually seeing a barn. And I further submit that the fact that the state as of seeing a barn is also correlated to his seeing something with such-and-such visible features is *not* what explains why the perceptual state disposes Henry to believe that there is a barn. So Henry is using HENRY-1 rather than HENRY-2.

Of course, whether I am right in thinking that Henry is using HENRY-1 depends on certain empirical claims I have made. If I am wrong about what explains Henry's disposition to believe that there is a barn upon seeing a barn, then it might turn out that Henry is in fact using HENRY-2 or perhaps some other method altogether. The crucial point for present purposes is that there is a real difference between Henry's using HENRY-1 and his using HENRY-2 which means that there is a real difference between his using a truth-conducive method which is too easy to misapply and his using a non-truth-conducive method which is not so easy to misapply. Mutatis mutandis for the other cases. Thus, all the conditions *i.* through *iv.* are needed to cover the various cases.

Let me address one more wrinkle before moving on. One might worry that an

account such as the one defended here which relies on conditionals that require consideration of a range of possible worlds for their evaluation will not be able to make sense of beliefs *de re* that are true by accident. For instance, if Henry sees a barn and judges that *that* is a barn, the proposition that he judges to be true is necessarily true given that barns are essentially barns. Because of this, it might seem that neither sensitivity nor safety can make sense of beliefs *de re* that are true by accident. If *p* is a necessary truth, it is trivially true that if not *p*, one would not believe that *p* and it is trivially true that if one believes that *p*, it would be true that *p*. Does my proposal face a similar challenge? It does not. Let me explain.

Let us consider two different cases. The cases differ in the method Henry is stipulated to be using. As just explained these stipulations incur further commitments as to the dispositions Henry has and what explains the presence of these dispositions but otherwise the cases are the same: Henry is driving through fake barn county and upon seeing an object on the road side, he forms the *de re* belief that that is a barn. The two methods are:

> **HENRY-3:** If you see a barn, judge that *that* is a barn!
>
> **HENRY-4:** If you see something with such and such visible features, judge that *that* is a barn!

Consider first the case in which Henry uses HENRY-3. The first thing to note is that the method does not by itself specify which proposition it is that one is to believe. Which proposition Henry ends up believing by using the method depends on what he is seeing. If he is seeing X, he will judge *de re* that X is a barn. If he is seeing Y, he will judge *de re* that Y is a barn. In my schematic presentation of methods as 'if C, judge that $P(w_s)$!', this is a case in which the value of $P(w_s)$ changes from world to

world and subject to subject. Now, this method is indeed truth-conducive but not because barns are essentially barns. It is truth-conducive because the satisfaction of the condition that one is seeing a barn entails that one is seeing a barn and hence that *that thing* one is looking at is a barn. In Goldman's fake barn county this method is easy to misapply. The belief one forms using this method has as its truth condition that the thing one is looking at in a given situation be a barn. In fake barn county it is very easy to form such a belief even though one is not looking at a barn. Conditions *iii.* and *iv.* fail in this case. Thus, beliefs *de re* about essential properties of objects formed by methods akin to HENRY-3 do not pose problems.

Consider next the case in which Henry uses HENRY-4. Like HENRY-3, the method by itself does not specify which proposition to believe. The proposition Henry ends up believing by using this method depends on the object he is seeing. HENRY-4 is not so easy to misapply in fake barn county. It is, however, not truth-conducive in fake barn county. It is not true in fake barn county that if one sees something with barn-like visible features, the thing one is looking at is a barn. So one can easily form a false belief by correctly applying HENRY-4: it can easily happen that one is seeing something with barn-like visible features even though that thing one is looking at is not a barn. So conditions *i.* and *ii.* fail in the case of HENRY-4.

Thus, neither HENRY-3 nor HENRY-4 pose a problem for my proposal. It should be clear how to handle other methods for forming *de re* beliefs. The reason why such methods do not pose a problem is that these methods do not require that one judge the same proposition to be true in all circumstances of its use: which proposition one ends up judging to be true using these methods depends on the

circumstances of application and this is why some of them come out as too easy to misapply and others as non-truth-conducive.[15] This, of course, leaves open the possibility that methods that require one to judge the same proposition to be true in all circumstances of its use pose a problem. In fact they do for the conditions *i.* through *iv.*

I noted above that *i.* and *ii.* cannot answer why the use of the method 'if a coin toss comes up heads, judge that Fermat's Last Theorem (FLT) is true!' leads to beliefs that are true by accident. The addition of the conditions *iii.* and *iv.* still cannot answer this question. So long as S is a normal human being who does not find it particularly difficult to determine whether a coin toss has come up heads, it will be true that if S believes that FLT is true via this method, the coin toss has come up heads; it will also be true that if the coin toss has not come up heads, S does not believe that FLT is true via this method. So the combination of *i.* through *iv.* does not capture the sense in which it would be an accident that S ends up with a true belief in a way inconsistent with knowledge if S were to use this method to judge whether FLT is true.

Here is a problem with the same roots. Consider again Luke's example above. He faces the choice of using a red scale or a blue scale one of which is unreliable. He tosses a coin and decides to rely on the blue scale. As a matter of fact, the blue scale is reliable. The method he uses for finding out his weight is: if the blue scale indicates that you weigh *n* pounds, judge that you weigh *n* pounds! This method

[15] To be fair, similar moves are available to a defender of either the sensitivity or the safety account so long as one is willing to avail oneself of belief forming methods and specifies them as I have. I see no principled reason why this should not be an option for either sensitivity or safety theory. My own objections to these views, as should be clear, do not depend on alleged difficulties handling *de re* beliefs.

is truth-conducive, i.e. conditions *i.* and *ii.* are satisfied. Conditions *iii.* and *iv.* are satisfied as well. Given that Luke uses this method, if he believes that he weighs *n* pounds, the blue scale will indicate that his weight is *n* pounds. And if the scale does not indicate that he weighs *n* pounds, he will not believe that he weighs *n* pounds. Yet there is a sense in which it is an accident that he forms true beliefs about his weight.[16]

One might perhaps hold that the problem here is that the method has been misidentified. Should we not say that Luke's method is something like

> **COMPLEX**: if the blue scale indicates that you weigh *n* pounds and you have decided through a coin toss that if the blue scale indicates you weigh *n* pounds you should believe that you weigh *n* pounds, then judge that you weigh *n* pounds!

This is not truth-conducive. It is not the case that if you don't weigh *n* pounds, either the coin toss came out differently or the scale does not say you weigh *n* pounds.

This is a kind of strategy suggested by Becker (2007, 103).[17] There is certainly something right about the idea here. It matters why one uses a given method and the way Luke comes to rely on the blue scale gets in the way of his having knowledge. However, this does not mean that we should think of the method that Luke

[16] It is important to keep in mind that our notions of luck/accident are distinct from our notion of objective probability. It is objectively quite probable that Luke will arrive at a true belief about his weight but he would still be lucky in getting it right.

[17] Becker's view is that knowledge is reliably formed sensitive true belief. It is a combination of process reliabilism with a version of the sensitivity requirement. According to Becker, the coin toss is part of the process by which he forms the beliefs about his weight. The process so understood is not a reliable belief forming process and this is why Luke does not know.

is using as somehow encompassing the way in which he came to rely on the blue scale. The natural way of expressing the situation is that the method Luke is using to figure out his weight is that of relying on the blue scale. The problem is that he came to rely on the blue scale through a coin toss. If our account can keep this rather natural distinction between the method one uses and the way in which one came to use it, that would be preferable. And the obvious thing to say here seems to be that, as already indicated earlier, it better not be an accident that one is using a truth-conducive method.[18]

2.3. It Must Not be an Accident That One Is Using a Truth-Conducive Method

The most powerful counterexamples to pure externalist theories of knowledge come from cases like Luke's (see, for example, BonJour's counterexamples to Armstrong's version of externalism (1985, 37–41)). It is simply not enough that one forms a belief using a method that is in fact truth-conducive. It is still not sufficient if it is no accident that one manages to conform to the requirements of a truth-conducive method. Something very crucial is missing when one leaves the explication of the non-accidentality condition at this: it matters why one is using a truth-conducive method. If it is an accident that one is using a truth-conducive

[18]Becker is aware that folding the coin toss into the belief forming process seems unnatural. His justification for nevertheless proceeding the way he does is that it gets cases like Luke's right and that it is allowed by his solution to the generality problem (2007, 103). It seems to me that he could simply add a clause like the one I discuss in the next section instead of using a non-standard way of individuating methods especially because it seems that his way of individuating methods is allowed but not mandated by his solution to the generality problem. However, given his commitment to the sensitivity requirement, my view is that his account so modified would still be objectionable.

method, then the resulting belief can only be true by accident.

Suppose Luke above *knows* that the blue scale is truth-conducive. In that case, we do not hesitate to attribute to him knowledge of his weight via reading the scale. Why is this? Our use of methods is often driven by beliefs; this is most evident in the case of reliance on equipments like scales, thermometers, tachometers, etc. If Luke correctly believes that the blue scale is truth-conducive, then this explains why he happens to be using a truth-conducive method. But it does not show that it is no accident that he is using a truth-conducive method. After all, he might have come to hold the belief through a guess or a coin toss. What is needed is that his 'choosing' the truth-conducive method be no accident. One way to satisfy this condition is by *knowing* that M is truth-conducive or, rather, by knowing the conditionals *i.* and *ii.* to be true and then basing one's use of M on this knowledge. In such a case, it would be no accident that what one believes to be a truth-conducive method is indeed truth-conducive (this is part of what it means to know), and hence it would be no accident that one is using a truth-conducive method.

So one way to satisfy the condition that it be no accident that one is using a truth-conducive method is to know that the method is truth-conducive and to base its use on this knowledge. Is this the only way of satisfying the requirement that it be no accident that one is using a truth-conducive method? If so, there would be a regress of a bad kind. In order to know that p via M, one would have to know that M is truth-conducive; i.e. knowing a proposition p_1 would require knowing at least one more proposition p_2, and this leads to the need of knowing infinitely many propositions in order to know just one proposition. This is not an acceptable consequence.

There indeed are methods that we use without basing our use of them on knowledge of their truth-conduciveness. Examples would be our use of perceptual methods, use of memory and basic logical/analytic reasoning. Let us call such methods 'basic' methods. Arguably, beliefs formed by such basic methods provide the starting points for our more elaborate body of knowledge. So if these 'basic' beliefs formed by using 'basic' methods do not amount to knowledge, there is no knowledge at all. How could it be that it is no accident that we are using basic methods which are truth-conducive?

When an agent uses a truth-conducive method based on her *knowledge* of its truth-conduciveness, the truth-conduciveness enters the explanation of why she uses that method. This is a crucial difference to a case where the use of a method is based on the agent's *belief* of its truth-conduciveness. What we want is that the truth-conduciveness of the method enter in the right way the explanation of our use of it.[19] Is it possible to provide such an explanation without appeal to the agent's knowledge of the truth-conduciveness of the method?

Let me suggest and sketch two possible explanations. One appeals to a causal story, in particular to evolution. Roughly, the argument would be that our ances-

[19]Reed (2000) proposes that one way in which a true belief fails to be knowledge is if it is justified accidentally. A belief has accidental justification, according to Reed, if "the explanation of why the belief is held is independent of what makes the source of that belief a reliable way of acquiring true beliefs" (2000,61). This is similar to what I am proposing here. However, it is not clear to me that all such cases are cases in which intuitively the agent is justified. On a more substantive note, what Reed means by calling a way of forming a belief reliable is that forming beliefs in that way is likely to lead to true beliefs. However, truth-conduciveness as I understand it is not reliability: truth-conduciveness only requires the truth of *i.* and *ii.* whereas reliability would also require the truth of *iii.* and *iv.* The importance of this point will be explained shortly.

tors who used truth-conducive belief forming methods were more likely to survive and reproduce than those who used non-truth-conducive belief forming methods. The use of truth-conducive belief forming methods conferred an adaptive advantage to our ancestors and this is why we, who are descendants of those who managed to survive and reproduce, use truth-conducive belief forming methods. Human beings evolved to use these methods precisely because these methods are truth-conducive and hence are fitness-enhancing. Since the truth-conduciveness of the methods enters the explanation of our use of them, we can say that it is no accident that we use them.[20]

The other possible explanation appeals to necessary conditions for thought. The methods in question here are methods for making judgments. Some such methods are truth-conducive but many of them are not. The question is why it is that the basic methods we use are truth-conducive. The answer rests on the idea that the ability to judge—i.e. our ability to think—depends on our using certain methods of judgment. For instance, if we did not use the method 'if X is a bachelor, judge that X is unmarried!', we would not understand what it is to be a bachelor. Methods the use of which are constitutive of being able to have thoughts are all truth-conducive.[21] Thus, the reason why the basic methods we use are truth-conducive is that only truth-conducive methods can enable us to think and hence enable us to use them—we cannot use methods of judgments unless we can make judgments.

[20] A view like Plantinga's (1993) proper functionism that appeals to an intelligent being that designed us to be cognizers is a close relative to the evolutionary story sketched here—the difference lies in the details of the causal story told: we might have an engineer, a breeder or natural selection that shapes our brains.

[21] Peacocke (1993) and Boghossian (2001) are two defenses of such a view of concept possession. Boghossian has since changed his view.

The truth-conduciveness of the basic methods enters the explanation of our use of them via the conditions of the ability to think at all.

It is vital here not to confuse the question whether a given method of judgment is truth-conducive with the question whether the use of a given method is likely to lead to true beliefs. The truth-conduciveness of the method used to form a belief does not guarantee the likely truth of the belief formed. A method 'if *C*, judge that *P(w_s)*!' is truth-conducive just in case *i.* if *C*, *P(w_s)* and *ii.* if not-*P(w_s)*, not-*C*. Whether the use of such a method is likely to lead to true beliefs depends on how easy/difficult it is for a given agent to do as the method requires. This was the point brought out by *iii.* and *iv.* So the view on concept possession just sketched is not committed to the *a priori* impossibility of massively false beliefs because the view is only committed to the position that the ability to think requires the use of truth-conducive methods of judgment and not to the position that it requires having true beliefs. A brain-in-vat is using truth-conducive basic methods: his problem is that he constantly misapplies these truth-conducive methods which leads to his having massively false beliefs.[22]

My own view is that a full explanation of why we use truth-conducive methods is a combination of the two views just sketched. A theory of the ability to think constrains what kinds of thinking creatures are possible at all: thought is only possible when the basic methods of judgment used are truth-conducive. But the actual existence of thinking creatures like us must be explained causally. The

[22] I do believe that considerations of what it takes to implement the use of this or that method rules out some skeptical scenarios (in the same way that content externalism may be used to rule out some skeptical hypotheses) but enough radical skeptical scenarios survive so that skepticism does not get ruled out *a priori*. I discuss elsewhere (?) in more detail what it takes to be implementing the use of epistemic methods.

theory of thought places constraints on what evolution can produce. Without this constraint, it is not at all clear that an evolutionary story can explain why we use truth-conducive belief forming methods. After all, what matters is that appropriate kinds of behavior are produced and appropriate behavior can very well be produced by beliefs, or even entire belief systems, that are widely off the mark—a point exploited by Plantinga in his argument against naturalism (Plantinga 1993, ch. 12). But given the constraint that thought is only possible through the use of truth-conducive methods, evolution can only produce thinking creatures, if it produces any, who use truth-conducive methods of judgments because thought is only possible through the use of truth-conducive methods.

Here is my fifth, and final, part of the non-accidentality condition:

v. the truth-conduciveness of M, i.e. the satisfaction of *i.* and *ii.*, explains why S is using the method M

This condition can explain what is wrong with using the method 'if a coin toss comes up heads, judge that Fermat's Last Theorem (FLT) is true!' When someone decides to use this method on a whim or somebody's say-so, the truth-conduciveness of this method plays no role in explaining why s/he is using this method. Similarly, a dogmatic person might be using the method 'if it seems to you that *p*, judge that *p*!' If *p* is true in a wide range of nearby possible worlds and if the person is such that it will seem to her that *p* in those nearby worlds in which *p* is true, this method will be both truth-conducive and hard to misapply. But the use of this method fails condition *v*. The truth-conduciveness of such a method generally plays no role in explaining why the method is being used.

I say 'generally' for it might be that for some instances of *p*, the use of something like 'if it seems to you that *p*, judge that *p*!' can lead to knowledge. I am thinking

of modal knowledge and the use of intuitions in philosophy. Whether we can take seriously our modal and philosophical intuitions as sources of knowledge rests on whether we can explain why for a certain range of propositions, only the true ones in that range are such that they seem to us to be true. In the absence of such an explanation, there will be a sense in which it is a mere accident that we are using truth-conducive methods in relying on our intuitions. I will not pretend to have such an explanation. Finding an account of how our modal knowledge works is a project of its own and this is not the place to pursue it. Suffice it to say that my conditions for non-accidentality can identify what the challenge is because my account of the non-accidentality condition is not restricted to empirical knowledge: *v.* gives a sense in which even true beliefs in necessary truths can be true by accident in a way that is incompatible with knowledge.[23]

3. Conclusion

Let me wrap up. The three parts of the non-accidentality condition were that given a true belief that p formed via a method M:

1. the method M used is truth-conducive
2. it is not an accident that one correctly applied M
3. it is not an accident that one is using a truth-conducive method

Examining these conditions one by one gives us that the following needs to be true if it is not an accident that an agent S correctly believes that p:

[23] In contrast, a theory that, for instance, treats safety as an explication of the non-accidentality condition cannot account for the sense in which beliefs in necessary truths can be true by accident.

it is not at all an accident that S's belief that *p* is true *if and only if* S used method *M* given by 'if *C*, judge that *P(w_s)*!' where *p* is *P(actual world$_s$)* such that

i. If *C*, *P(w_s)*

ii. If not-*P(w_s)*, not-*C*

iii. If S believes that *P(w_s)* via *M*, *C*

iv. If not-*C*, S does not believe that *P(w_s)* via *M*

v. the truth-conduciveness of M, i.e. the satisfaction of *i.* and *ii.*, explains why S is using the method M

Note that it is tempting to conclude from *i.* and *iii.*: if S believes that *P(w_s)*, *P(w_s)*; and from *ii.* and *iv.*: if not-*P(w_s)*, S does not believe that *P(w_s)*. The first would be the crucial conditional for safety, the second the crucial conditional for sensitivity. But, as I have already pointed out, such inferences are not valid given the kinds of conditionals involved. Knowledge requires neither sensitivity nor safety.

Nevertheless, given *i.* through *iv.* for particular *C* and *P(w_s)*, it might very well turn out that the safety and sensitivity conditionals are both true and this might explain why both the sensitivity and safety conditionals look at first glance rather tempting. But we must not forget that the pairs *i.*, *ii.* and *iii.*, *iv.* explicate quite different aspects involved in the non-accidentality condition.

We can think of process reliabilism, sensitivity and safety theories as concentrating on various aspects brought out by *i.* through *iv.*: is the method used any good for answering the question whether *p* and is one in a situation such that one is unlikely to end up with false beliefs? They suffer, however, from failure to distinguish different sources of accidentality. Internalist criticisms of pure exter-

nalist theories of knowledge concentrate on *v.*: it better not be an accident that one is using a truth-conducive method. Of course, epistemological discussions are much muddier than this would suggest but reflection on how to satisfy the non-accidentality condition does provide some template for mapping out what various epistemological theories of knowledge are contributing to an overall understanding of knowledge.

One final remark. Unger's analysis of knowledge says that it is necessary and *sufficient* for S's knowing that *p* that it be not at all an accident that S's belief that *p* is true. I noted that one might hesitate about the sufficiency claim. But given the above discussion, it seems that the sufficiency claim is indeed correct. What more could be needed apart from the non-accidentality condition? One might think that the use of the method must also be *rational* in addition to being non-accidental but to the extent that such a demand is reasonable, it is covered by *v.* since in many cases *v.* is satisfied by the agent's knowing that the method she is using is truth-conducive (surely, if the agent knows the method to be truth-conducive, she is being quite rational in using it) and in those cases where the condition is satisfied by the concept constituting character of the method used, it is hard to see how the use of the method can fail to be rational if the use of it is required for the ability to think in the first place. One might make the unreasonable demand that one must always know in advance, as it were, that the method one is using is truth-conducive. But such a demand seems to amount to a particular view on how condition *v.* is to be satisfied and not to a distinct further condition on knowledge. So I conclude:

S knows that *p* if and only if it is not at all an accident that S's belief that *p* is

true.[24]

References

Becker, K. 2007. *Epistemology modalized.* New York: Routledge.

Boghossian, P. 2001. How are objective epistemic reasons possible? *Philosophical Studies* 106.

BonJour, L. 1985. *The structure of empirical knowledge.* Cambridge, MA: Harvard University Press.

Comesaña, J. 2005. Unsafe knowledge. *Synthese* 146 (3): 395–404.

Davies, M. 2004. Epistemic entitlement, warrant transmission and easy knowledge. *Aristotelian Society Supplementary Volume* 78 (1): 213–245.

DeRose, K. 1995. Solving the skeptical problem. *The Philosophical Review* 104 (1): 1–52.

Dretske, F. I. 1970. Epistemic operators. *The Journal of Philosophy* 67 (24): 1007–1023.

Foley, R. 1984. Epistemic luck and the purely epistemic. *American Philosophical Quarterly* 21 (2): 113–124.

———. 1987. *The theory of epistemic rationality.* Cambridge, MA: Harvard University Press.

[24]I would like to thank Yuval Avnur, Nico Sillins and Peter Unger for helpful comments on earlier versions of this paper. I am also indebted to anonymous referees for many insightful comments.

Gettier, E. L. 1963. Is justified true belief knowledge? *Analysis* 23 (6): 121–123.

Goldman, A. 1979. What is justified belief? In *Justification and knowledge*, ed. G. Pappas. Dordrecht: Reidel.

———. 1976. Discrimination and perceptual knowledge. *The Journal of Philosophy* 73 (20): 771–791.

Neta, R., and G. Rohrbaugh. 2004. Luminosity and the safety of knowledge. *Pacific Philosophical Quarterly* 85 (4): 396–406.

Nozick, R. 1981. *Philosophical explanations.* Cambridge, MA: Harvard University Press.

Peacocke, C. 1993. Proof and truth. In *Reality, representation and projection*, eds. J. Haldane and C. Wright. New York: Oxford University Press.

Plantinga, A. 1993. *Warrant and proper function.* New York: Oxford University Press.

Pritchard, D. 2005. *Epistemic luck.* Oxford: Oxford University Press.

Pryor, J. 2004. What's wrong with Moore's argument? *Philosophical Issues* 14 (1): 349–378.

Reed, B. 2000. Accidental truth and accidental justification. *The Philosophical Quarterly* 50 (198): 57–67.

Russell, B. 1997. *The problems of philosophy.* Oxford: Oxford University Press.

Sosa, E. 1999. How to defeat opposition to Moore. *Philosophical Perspectives* 13:141.

Unger, P. 1968. An analysis of factual knowledge. *The Journal of Philosophy* 65 (6): 157–170.

Warfield, T. A. 2004. When epistemic closure does and does not fail: a lesson from the history of epistemology. *Analysis* 64 (281): 35–41.

Williamson, T. 2000. *Knowledge and its limits*. Oxford: Oxford University Press.

Wright, C. 2002. (Anti-)sceptics simple and subtle: G.E. Moore and John McDowell. *Philosophy and Phenomenological Research* 65 (2): 330–348.

Yamada, M. 2010. Rule-following: a pedestrian approach. *Philosophy and Phenomenological Research* 80 (2): 283–311.