Spring 5-17-2020

Ascribing Knowledge: An Approved Reliable Methods Account

Xiaoyu Su
Dickinson College

Ascribing Knowledge: An Approved Reliable Methods Account

by

Xiaoyu Su

Submitted in partial fulfillment of Honors Requirements

for the Department of Philosophy

Dr. Susan Feldman, Advisor

Dr. Chauncey Maher, Reader

Dr. Amy McKiernan, Reader

May 12, 2016

“Knowledge is in the end based on acknowledgement. ”

Ludwig Wittgenstein, On Certainty, No. 378

“Every explanation is after all a hypothesis.”

Luwig Wittgenstein, Remarks on Frazer’s Golden Bough

Table of Contents

Introduction

Epistemology concerns questions about knowledge and primarily propositional knowledge. A primitive version of the question is the classical “what is knowledge?”. Questions like this are oftentimes not very interesting because it does not indicate what we wish to unravel with regards to knowledge. We then realize that it is better to consider knowledge as a kind of property or some sort of achievement that belongs to certain subjects, for knowledge, if isolated from the cognizers, is nothing more than a collection of true statements. Epistemologists have been looking for the conditions where a knowledge attribution to a subject is true. Generally, the question we are asking is, for a subject S and a proposition P, what is the truth condition for “S knows that P”? From there, epistemology has been dealing with the problems of skepticism and devising various accounts to account for carefully designed scenarios that rarely take place in actual life (e.g. the Gettier cases). While thinking about these intricate cases were interesting, it left me wondering, sometimes, what brings me into this discussion in the first place?

Consider the situation of a group of migrants fleeing a pandemic and famine. Certainty becomes important and worth pursuing when their survival depends on it. Imagine one of the group was sent to scout for supplies. The scout had better make sure his observational report (think of it as a proposition) was trustworthy. The rest of the group would hold him responsible for his report, for the consequences of him being wrong would be disturbingly negative, while his being right might just save the whole group. It is crucial for the group to evaluate his report, and group members would then plan their decisions based on the evaluation. Thus attributing/not attributing knowledge to the scout is essentially an act

(instead of a mere proposition) with purposes and consequences. Knowledge attributions lead us to further actions and provide us grounds to believe that the outcome of those further actions is promising. Indeed, it is these acts of knowledge attribution that makes knowledge an important aspect of our communities and so an intriguing philosophical inquiry. Hence, in order to understand the nature of these acts, the key questions I should be pursuing are "What are we doing when we are attributing knowledge? ", "How does knowledge attribution work within a community?", and "How is the language game concerning knowledge attribution supposed to be played?".[1] To answer the question, in what follows, I am going to focus on the assertibility conditions instead of the truth condition for 'S knows that p', as I believe this sets us on the right track to understand knowledge attributions as acts and how these acts are embedded in human communities.

Assertibility conditions for a statement basically translates to whether a statement is appropriate to be made in certain contexts.[2]The appropriateness is not to be understood in terms of something that limits the number of statements people should be saying in a social environment, rather, we say that a statement is appropriate to assert if the statement is used correctly, or that the use of the statement should be approved in the given situation.[3] Why is this a better approach than seeking for the truth condition of knowledge attribution? Because

[1] "language game" is an idea developed by Wittgenstein in *Philosophical Investigation* (1953). This idea considers language as a kind of activity that involves a set of rules, which resembles the games we've played.

[2] This idea is originated from Brandom's 1976 paper "Truth and assertibility".

[3] Here is one way to think of how it differs from the truth condition. Let's say that, in reality, oranges are different from all the other entities in the world in that that an orange pops out of existence when no human minds are attending to it. If this is the case, we might suggest that a lot of propositions regarding oranges are false. However, if this situation is unknown to us, our normal usage of language concerning orange remains functional. As long as this situation is unrevealed, we should still be approving and denying statements concerning orange as we normally did.

rather than considering whether knowledge attributions are true or false, I think it is more fruitful to explore what makes them appropriate or not. Human communities always work with a restricted set of information that prevents them from being the ultimate judges of truth or falsity (when it comes to empirical propositions). That is, every such statement we've made might potentially be falsified. Thus our empirical beliefs including those of science were and are always subject to constant revision and updates based on new information. The talk of truth/falsity is unable to capture all the reasonable moves we have made in the game of knowledge attribution.

Here is one example of such moves. Consider the fact that we rely on devices to secure beliefs. We have thermometers to tell us about the temperature; a click on "refresh" informs us whether we have new emails; fingerprint matching systems help us make sure somebody had been to some places. The devices we use in these cases are by no means perfect in their functionality and the way we secure our beliefs using these devices are non-flawless heuristics. Thus if we perform countless trials using these belief heuristics, we are eventually going to encounter a failure on a certain trial, even if the evidence we collected this time is of no fundamental difference than that of other trials (or at least it appears to us this way). Hence if we attribute knowledge here, the attribution would be false. However, we nevertheless feel like it is reasonable to do so, for we construe this episode in the same way we construe other ones. The nature of thinking about the assertibility condition is to not focus too much on one single failure, but to look at the larger picture, namely, to look at all the related episodes as a group. What we should observe here is that we are in need of systematic

approaches to attribute knowledge that operate effectively, though not perfectly, in our communities.

For example, watches take care of beliefs concerning time in normal settings. If we simply attribute knowledge if the cognizer takes a responsible look at the watch, things operate smoothly almost all of the time. It is mostly accurate and favorably fast. For the occasional meltdowns (if the meltdowns are not occasional, we wouldn't be relying on those watches in the first place), because they are occasional, it doesn't cost too much just to accept the failures and ignore them. And if we want to eliminate these meltdowns using a new heuristic, it can cost too much and so it is practically impossible. This watch --> time heuristic remains our status quo since alternative heuristics either perform poorly or are economically inefficient. If we think about our community where its members attempt to deal with types of beliefs through systems of cognitive heuristics, it is then evident why talking about the assertibility condition for knowledge attribution is more fruitful. It essentially describes the shape of knowledge attributions by considering the purposes communities wish to realize. In this case, we have simple and reliable ways of securing beliefs and attributing beliefs concerning time that allows further related beliefs, actions, and coordinations of the community to take place. In order to move forward satisfactorily, we must accept the fact that our means of securing beliefs might occasionally fail.

If we look into how the way we conceptualize the world evolves over time, we do see that we were constantly substituting some beliefs with others while both of them are sets of appropriately attributed beliefs. Newton's second law of motion $F = m \cdot a$ was appropriately counted as knowledge based on the numerous experiments conducted back then, until it was

figured out that this formula does not apply to close-to-speed-of-light objects. We now only consider the restricted version of the formula to be approximately true. But it is still possible that hidden variables would overthrow the current conception (Einstein's) and yield a new one. However, even though the original formula is not universally correct and the new formula has an undeniable potential to be falsified, the sense of appropriateness in them is the actual appeal of the language of knowledge. At each stage in the human history, the important thing was to have a communal commitment to reasonable beliefs (or to take reasonable beliefs to be true) on the basis of our effort. If we only consider the truth condition of knowledge attribution, the only thing we would say is that attributing knowledge to Newton in this case is false, while the historical position of Newton's second law of motion in human knowledge library is left unexplained. It is thus urgent to figure out the conditions for asserting 'S knows that P' to understand how our knowledge system works.

Skeptics are right in pointing out that we do not have perfect heuristics for arriving at beliefs. But language never asked for such luxury and so the skepticism concerning knowledge is beside the point. Conversations about knowledge should not aim to provide us with a maximum level of certainty human beings dreamed of, but to support us in achieving things that cannot possibly be achieved without such a language deployment (e.g. report the bear you see in the woods before it becomes a threat to the community). Skepticism thus only educates us on underdetermination and the impossibilty of the God's eye point of view, but it does not satisfy us in terms of how knowledge as a concept operates in our community. Much traditional epistemology focuses on the truth conditions for "S knows that p', but a move to assertbility conditions --- the circumstances under which we appropriately attribute

knowledge of P to S allows us to avoid constant entanglement with skepticism. Assertibility conditions ask us to think about the purposes and circumstances of knowledge attribution. Skepticism is something a community must ignore to realize those purposes.

In this thesis, I consider knowledge attributions as connected to methods generating beliefs. The main argument of is that knowledge attributions are appropriate when the cognizer arrives at the beliefs through methods that are deemed reliable by the community in certain contexts. I only intend this argument to help us aim at the center of the target, as the whole issue is too complicated and interesting for any single statement to cover. My thesis consists of three sections. In the first section, I present Goldman's account of reliabilism as well as McDowell's and Brandom's critique and discussion of Goldman's theory. I defend reliability as the key to understanding knowledge attribution in terms of its tenability and advantages. Construing knowledge as a product of reliable methods provides us with a framework in which both the insights of *externalism* and *internalism* can be incorporated. In the second section, I provide a detailed model (a T/F ratio test) explaining the way we (ideally) measure reliability. In the third section, I illustrate a way of understanding reliable methods (approved-list reliabilism) that avoids the objections and the concerns arising about reliabilism in general. I argue that we should consider the set of reliable methods as a list or a toolbox developed and assessed by a community and subject to change over time, rather than an unmovable ontological kind. Then I consider the rationale or epistemic standards for managing this list and conclude that it is impossible to make explicit a clear set of criteria that accounts for what gets to be on the list.

Section 1:Why Reliabilism?

Knowledge is traditionally defined as justified true belief (the JTB account). Philosophers almost unanimously agree that only true statements can potentially be counted as knowledge. But true belief itself is not sufficient for knowledge. Under most JTB's interpretations, what distinguishes knowledge from mere true beliefs is the subjects' appropriate justification for holding the true belief. We wouldn't say S knows that P if S happened to claim a true belief P. On the other hand, if S provides solid reasons or cites good evidence, we are far more willing to say that S knows that P. As an example, if S claims "there is no one in the classroom", we usually require S to tell us why S believes so. In this case, guessing would be bad reasoning while "checking with one's visual perception" would be a good justification. Whether we attribute knowledge to S is determined by how good S's justification is or how well S justifies herself. S's belief P is justified on the basis of S's direct or indirect evidence, reasons, or experiences. These traditional JTB accounts of knowledge are characterized as forms of *internalism*, since they emphasize the salient effort the subject made in order to get to the belief; under these accounts, the knowledge status of the subject is determined and only determined by the subject's mental states and justification on the basis of these mental states.

But what counts as being justified is itself a complicated issue. A traditional criterion says that a belief P is justified as long as the justification establishes the truth of P. (Williams 2001, 52) This seemingly promising idea faces the problem of skepticism. Cartesian skepticism claims that there are no mental states that perfectly establish the state of affairs as it is. In other words, our personal experience always underdetermines the reality. There might be an infinite number of realities that would produce the same collection of experience.

Descartes' evil genius problem says that our mental states can be the same even if, instead of being in a normal world, we are captured in the evil genius's illusions. Thus our mental episodes do not guarantee the world being what we typically infer from those mental states, which leaves us an unfortunate dilemma where knowledge seems impossible. The famous Brain-in-a-Vat thought experiment attempts to do something similar. How do we know that the experience we took to be real wasn't the product of advanced brain simulation? We can be skeptical even without skepticism of these extreme sorts. Think about our memory capacity. Plenty of our justification for beliefs is relying on the ability of memory retrieval. However, we do experience cases where we held beliefs based on memory but those beliefs turned out to be false. Since the next time we exert our memories seems to be no different from our previous exertions, and we cannot rule out the possibility that our memories do not work, we cannot safely guarantee that the fact is secured from our memory.

We don't want to say that knowledge is impossible. After all, there is something we are doing and achieving by engaging in all these conversations about knowledge. Such a negative view does not provide us any insight into the role the concept of knowledge plays in our linguistic behavior and community activities. This moves us to fallibilist accounts of knowledge since we can no longer explain the status of being justified by appealing to "justification that perfectly secures the fact". (Williams 2001, 41) Fallibilism is the idea that knowledge can be attributed without absolute certainty. In other words, knowledge attribution can and should be made even if the evidence from the cognizer does not guarantee the belief being true. We are left with the intuition that memory should be a solid source of knowledge, on the one hand, and that memory is not perfect on the other, and they seem to be

incompatible with each other. But do observe that even though memory is not perfect, our intuition says that memory is fairly reliable in general and its ability is trustworthy. Also, we are more likely to trust a teenager than an elder who suffers from Alzheimer's in terms of memory and this discrepancy is precisely a result of the two capacities' different levels of reliability. That is, the teenager can be expected to arrive at mostly true beliefs through memory retrieval, while the elder can be expected to generate falsehoods way more frequently relying solely on memory. So maybe the way we attribute knowledge is not by looking at how experience validates states of affairs, but by looking at how reliable we are at generating true beliefs.

Alvin Goldman was an early proponent of a theory of knowledge built upon the idea of reliability. In "What is Justified Belief?", he attempted to explain justifiedness not in terms of traditional justification appealing to the mental states of the cognizer, but in terms of the reliability of a belief-forming process. (Goldman 1979) In particular, the justified beliefs are just the ones that are formed through a process that has a high true/false ratio in the past. In contrast to *internalism*, his view is a form of *externalism* (Williams 2001, 31) because something other than the cognizer's current mental states determines her knowledge status.

Goldman observed that plenty of times when we are unjustified, our beliefs are caused by some faulty belief-forming processes. (Goldman 1979, 95) Confused or logically fallacious reasoning belongs to this category. A quick glance can also be a process of this sort. Sometimes a glance made me believe that the black little dots on my table were bugs. But with a closer look, it turned out that they were pieces of dirt. Thus if we find that the cognizer arrived at his belief based on processes like these, we would mark his belief as unjustified.

On the contrary, justified beliefs can be seen as results of standard or good belief-forming processes. In mathematics, if we correctly apply valid deductive reasoning given the assumptions, we would always arrive at justified beliefs. Goldman then adopted the idea of reliability to address this issue. Reliable processes yield justified beliefs while unreliable processes yield unjustified beliefs. Simply put, Goldman's main thesis is that some true belief counts as knowledge if and only if the cognizer forms that belief through a reliable process. According to Goldman, classical reliable processes include good perceptual processes, memory, good reasoning, whereas wishful thinking, reliance on emotional attachment and similar sorts of things are the unreliable ones. Reliability is a matter of degree. Think about the fact that our confidence in our beliefs caused by visual observation dwindles as the distance of observation increases. Close examination is more reliable than a casual long-distance glance. But what exactly measures the reliability of these processes? Goldman's answer is through historical records. A process's reliability is determined by its historical ability to produce true beliefs. A reliable process is just a process that would yield a significant more amount of truth than falsity based on having produced a significantly high true/false ratio in the past. The reason we have this impression that careful examination is better than a casual glance is that, historically, the former produced true beliefs at a higher rate than did the latter. Based on this insight, Goldman marked his account of justified beliefs as a "Historical or Genetic" theory as opposed to the more common "Current Time-Slice" theories (e.g. the traditional JTB accounts) which made justifiedness a function of the current status of the cognizer. ((Goldman 1979, 98)

In their 1995 discussion of Goldman's externalism, McDowell and Brandom characterized Goldman's account as a kind of *pure externalism* in the sense that justification or reason is completely out of the picture. In "Knowledge and the Internal" (McDowell 1995) by McDowell and "Knowledge and the social articulation of the space of reasons" (Brandom 1995) by Brandom, both of these authors agree that externalism of this extreme sort is problematic. A counterexample is that such a theory would allow a thermometer or a parrot to have propositional knowledge which we would consider as something that can only be possessed by humans. (Brandom 1995, 896-897) Later we will see a version of *hybrid externalism* that incorporates external factors as well as the internal ones.

Brandom argues that externalism of this extreme sort does not satisfy the belief requirement in having knowledge. A parrot can respond with the utterance "this is a green ball" upon being presented with a green ball reliably. But the parrot does not count as a knower particularly because it does not actually hold a belief. Following Sellars, Brandom contends that being able to take a position in the space of reasons (one must be responsive to what one would commit himself to in making a claim, what he would be making himself responsible for, what is the evidence for/against it, what is compatible/incompatible with it, so on and so forth) is essential to having concepts and beliefs. In order for one to form a belief about something being green, not only does one have to respond differentially and reliably to green things and non-green things, but she also has to be applying the concept "green", that is, she has to know the inferential role "green" plays. For instance, she has to know how "green" is different from many other concepts, what the evidence for it is and what something being green means. We cannot say that a parrot who reliably utters "that's green" to green things or

a thermometer that reliably reports the temperature of the room are expressing beliefs because a parrot or a thermometer cannot understand concepts let alone apply them. Brandom contrasts a 2-month-old child and a 5-year-old child uttering the same sentence. While the former simply makes noises, the latter began "master the inferential articulation of the space of reasons". (Brandom 1995, 898) In other words, he started to form a general idea of what he is talking about. He can respond to questions such as "What do you mean by the word 'X'?" and "Why do you believe this is the case?". Brandom acknowledges that reliabilism is right in suggesting that we are able to have knowledge even when we are not able to defend ourselves. One classic example is identifying faces. It is very hard for us to explain how we manage to differentiate one face from another, as we do not rely on linguistically explicit approaches. But even though we are unable to provide justification, as long as we can differentiate faces reliably or at a high success rate, we can be ascribed knowledge without a problem. However, not being able to provide reasons does not entail that reason need not play a part in the process of acquiring knowledge at all. The difference between parrots and humans is that only humans can provide further explanation for his utterance (e.g. explaining the concepts, rewording the statement). Being able to respond with true statements reliably is not good enough when it comes to knowledge attribution.

Given that reasons must be involved, how about making the space of reasons an extra requirement for reliabilism? The attractiveness of reliability and the indispensability of reasons and responsibility leave us with the choice of combining them together. Brandom and McDowell call this "the hybrid view". If we count the availability of the space of reasons as a

necessary condition for holding a genuine belief, wouldn't we just preclude the "fake knowers" from joining the knowledge game?

McDowell believes that this view is still problematic. In particular, McDowell specifies two ways of understanding the idea of reliabilty and deemed neither satisfactory. Firstly, McDowell argues that if consideration of reliability is beyond the reach of individual's reasons, then there would be no connection between justification and reliability. (882) This is to say that if the notion of reliability only accounts for the objective reliability of different cognitive processes, reliability is then disconnected from our understanding of how reliable a certain cognitive procedure is. Our grasp of the reliability of, say, visual perception, is through our impression of the historical records of the performance of such cognitive process. Our impression of these would vary from time to time and whether our beliefs are justified depends on our impression of the reliability of the process. Thus if reliability is to be understood as something objective and invariant, it cannot possiblybe incoporated into practices of knowledge attribution. Brandom has no problem agreeing with McDowell here.

If, however, we incorporate assessments of reliability into the standings of the space of reasons, that is, we make it a product of our empirical life, for McDowell, it doesn't work either. Since reliability only tells you a process is more likely to produce truths rather than falsity, it does not guarantee "a cognitive purchase on an objective fact". (McDowell 882) What McDowell means is that forming beliefs reliably does not entail that the world must correspond to the content of the belief. Since there is a chance that the reliable process would produce false beliefs, I cannot fully determine whether I am lucky and the reliable process offers me a truth or I am unlucky and pick up a falsity. However, assessments of reliability

have to be built on top of the historical record in which every proposition is marked as either true or false. So let's say I adopt a cognitive process the reliability of which is yet to be assessed. I wouldn't be able to mark my beliefs based on this process as true or false because I am yet to determine its reliability. But without the record, I wouldn't have the resources to make assessments of reliability. This leaves reliability an untenable notion.

Brandom denies that this untenability exists. The way out, for Brandom, is by urging us to consider knowledge, not as something one has to achieve all on one's own or individually, but a social status that is constituted both by the cognizer and the knowledge attributors. For Brandom, to view knowledge as standings in the space of reason is to view assertions as commitments and entitlements, which are social statuses we offer each other in the game of "giving and asking for reasons". In understanding knowledge as a social product, the key is to make it clear what happens when one attributes knowledge to the cognizer who makes a commitment. Brandom suggests it involves essentially three components: attributing a commitment, attributing an entitlement, and undertaking that commitment. The first one suggests that the attributor would hold the subject responsible for her claim and so it is appropriate to ask for her evidence, to ask what she suggests to do and so on. The second thing suggests that the attributor takes the process through which the subject gets to the claim as something reliable. The third process suggests that in attributing knowledge, the attributor must herself be making the same commitment. This should intuitively make sense. If Bob believes the marker is out of ink and we think he knows it, then we ourselves must believe the same thing. (undertaking a commitment) Also, we must believe that Bob takes his belief

to be true. (attribute a commitment) Finally, we must believe that Bob arrived at his belief properly. (attribute an entitlement)

Understanding knowledge as bound up with interpersonal relations allows us to see that assessments of reliability and consideration of truths are carried out by the knowledge attributors rather than the cognizer, which then free the cognizer from the responsibility of assessing her reliability. Consider how we credit 5-year-old children with observational knowledge. The children themselves might not be capable of conceptualizing their own reliability. They are unaware of it when forming beliefs and are unable to provide it as a reason for their beliefs. But there is no problem for us to ascribe knowledge to them when they continued to demonstrate their ability to do observational reports. In fact, no one can possibly be a knower on her own without the assessments of other community members. As Brandom suggests, "it is the different perspectives provided by different sets of commitments that make it possible to triangulate on objective states of affairs." (907) The reality of color, odor, sound, etc. are formed on the basis of the convergence of perspectives. We can thus view our language games as a product of this triangulated reality and we serve as supervisors of each other to check how well we converge with other people. Any individual might be unaware of his cognitive system being different from the rest of the community. But if someone perceives the world differently in certain categories, she cannot be deemed a knower in those categories. More precisely put, she wouldn't be able to join the language game of knowledge attribution other community members are playing. Since it is evident that she herself wouldn't be able to identify whether her observation converges with those of other people, this process can only be done when the community members are exchanging

their beliefs with her. Hence an individual is unable to determine her own reliability even if she realized her full capacity or be fully responsible. I might not realize my defects that affect my ability to have knowledge in certain categories even if I try my best. I might wake up the other day and have a completely isolated auditory perspective that goes against the perspectives of others. Other people are not only important but essential to the assessments of my reliability. There is no untenability in assessments of reliability according to truths essentially because truths are not separated from the perspective of the community members.

McDowell further contends that if two subjects are alike in terms of their reliability, but one has knowledge while the other doesn't, what makes the difference cannot be anything internal to them (since the reliability does not guarantee the world being so) and so it must be a favor the world does for the former.[4] Reliabilism is thus problematic because it considers knowledge as something close to mere true beliefs. Reliability only moves one closer to factiveness; it is luck that brings one to the truth, hence the hybrid view misses something important about knowledge.

Brandom did not address this issue. But I have an answer for this concern. McDowell's argument should remind us of the Gettier cases. (Gettier 1963) Here is one example of Gettier cases If S says that the final score of a basketball game is 90 to 100 because of a creditable website reports so, it is very plausible to say that S is justified in saying so (call this the website case). But in reality, the website's report is a random mistake and it is a pure coincidence that the score of the game was exactly 90 to 100. Our intuition says that S does not know the score,

[4] Consider A and B who are equally reliable in terms of identifying a kind of object O. SInce their reliability is not perfect, they might occasionally hold false beliefs. McDowell thinks that if this time A forms a true belief about O while B makes a false claim, since their reliability is the same, the situation can only be explained by luck, which he takes to be unsatisfactory.

for it is luck that brings S to the truth. Observe that in both McDowell's argument and the Gettier cases, we seem to have some reliable processes (i.e. the reliability of the individual & the reliability of the website) that generate us true beliefs out of luck. But there are two different senses of luck concerning conversations of knowledge attribution. One is unacceptable but the other is not. If I randomly form a belief P and it is luck that offers me a true statement, it is unacceptable and thus I do not know P. However, if I formed a belief based on a reliable process and it is luck that prevents me from the slim chance of being wrong, then knowledge attribution wouldn't slip away. Here we see why it is important to distinguish between the assertibility condition from the truth condition of 'S knows that P'. Given that S is informed by a creditable website, it is appropriate for us to say S knows that P, even though it does not guarantee the truth of "S knows that P" (since weird situations might take place). Thus we can have an appropriate attribution but a false one. It is correct for us to attribute knowledge since S's belief is formed based on a trustworthy website. However, if it is disclosed that weird situations took place, we should then reevaluate S's knowledge status. This model of language game works only because the chance of having a malfunctioning reliable process or a Gettier case is small enough that the risk of a mistaken attribution can be ignored. And this is the only model that works because under the logically compatible possibility of skeptical scenarios, it only takes a little imagination to find out that we are never in a position to fully determine the truth value of a proposition P. What we as a community do is to send out knowledge attributions as efficiently and economically as possible and take them back when things go against our conception. Taking attendance by calling students' names allows the professor to know who is present and who is not in a time-efficient manner which further leads to the execution of certain

grading policies. But if we are worried about the slim chance of someone in class being a hard-to-distinguish impersonator, either we develop a stricter method that isn't economical and thus isn't realistic or we are preventing ourselves from drawing any conclusions. Either way, our knowledge status concerning students' presence is severely threatened and our actions thereafter are impeded. But if we commit to the idea that we need to ignore those kinds of improbable scenarios and make adjustments if those scenarios are known to take place, then things would work just fine. The way language games concerning knowledge works is that we sacrifice a little security for high efficiency and executability. The benefits of attributing knowledge in this case far exceed the risk of falling into one of the "unlucky" scenarios. And only when we choose to ignore the improbable events are knowledge attributions possible.[5] This is why the "luck" McDowell worried about is not a problem and so his concern was misplaced. This model would dissolve our confusion towards Gettier cases as well. In the website case, it is absurd to abandon all assertions of knowledge attribution just because Gettier cases might strike. But our assessments of knowledge status should vary according to the information that has been disclosed to us. Hence when we only heard that S viewed the website, we should say S has knowledge, and when we have the full information, we should say S does not.

Reliabilism of this sort can handle the cases in which the ability to justify does not seem to be necessary (non-inferential, perceptual knowledge, such as facial recognition or chicken sexing). Indeed, we wouldn't collect concrete and explicit evidence when we adopt some of the reliable processes. Brandom points out nicely that chicken-sexers do not have a specific

[5] This idea is originated from David Lewis's "Elusive Knowledge". (1996)

piece of evidence or mental episode that allows them to distinguish male chicken from female ones. (Brandom 1998, 375) Hence these people wouldn't be able to explicitly defend their beliefs or provide a clear line of justification as to how they come to form their beliefs about the chicken (even though they still need to be able to apply reasons in general). Or we can think of how we recognize people we are familiar with. We would say something like "I just know" or "I don't know how, but I do recognize that this is my brother".[6] In these cases, ascribing knowledge would usually be associated with how frequently we get things right (i.e. reliability) rather than the reasons we provide. But note that these are the minority of cases. More often than not, there does seem to be crucial explicit rational support and it is normative for cognizers to tell their stories to knowledge evaluators, otherwise it might not even be possible for the latter to make a judgement.[7] If the idea of reliability is to satisfy all the sorts of knowledge attribution we have been doing, we cannot get away with the inability to justify. What motivates the JTB account of knowledge is exactly how we adopt justification, reasoning, and evidence in language games concerning knowledge.

A good reliabilist theory of knowledge shouldn't be opposed to these concepts, instead, it should provide a larger framework in which these concepts can be accounted for. I believe this is not only possible but also necessary. There are two argument for this. (1) Internal mental states and justification, as important as they are, cannot themselves determine the knowledge

[6] Here is an interesting example from my life. Usually, the left and right earbuds are built differently. People can tell which one is which just by touching them. But it extremely hard to verbalize that kind of experience. Moreover, the verbalization wouldn't be the things people look for when attributing knowledge in these situations. Instead, we need cognizers to repeatedly be able to tell left or right just by touching the earbuds.

[7] This is definitely the case for inferential knowledge. For non-inferential knowledge, consider the situations where the knowledge evaluators cannot verify the beliefs themselves. For example, if someone came back from 100 miles away and report what was happening there, we can't possibly verify his beliefs. In situations like this, we demand justification.

status of certain beliefs. External factors beyond the experience of the subject play an indispensable role. (2) The evaluation of knowledge status requires the evaluation of the reasons offered as justification. Justification does not itself explain its quality, or how strong it is, or how frequently it can lead to true beliefs. Thus when it comes to measuring and assessing the strength of justification, reliability or ideas of its sort must invariably be invoked. I will treat each argument in turn.

First, justification comes from mental states available for a subject S. My claim is that mental states themselves are not sufficient in determining the knowledge status of S. This would be true if S has the same mental states as others in certain episodes and yet cannot be said to possess knowledge while others can. Consider the conceptually possible scenario in which one day S's neural system suddenly started to transform colors in her color vision into random colors. Before this happened, S was a standard observer who learned to use the color language like everyone else did. After this randomization pathology, what was originally green to S can present as red, yellow, or any other color. Suppose by luck S is appeared to greenly in one episode and correctly reported "this is green" but was appeared to redly to the same thing in the next episode with a similar environment. If others are aware of S's weird symptom, they wouldn't ascribe color knowledge based on S's vision to S. But just in that case where S reported green correctly, S supposedly had the same mental state and provided identical justification as other people do (and as S used to before the pathology set in) yet it does not make S a knower. The takeaway is that we say we know what is green not only by being presented with greenness upon the presence of green things, but by being able to achieve that in a consistent manner. It follows that a single episode of mental states is insufficient in terms

of ascribing knowledge. We want to make sure that it is no accident that S collects certain evidence upon certain states of affairs. But consistency is not to be found in any single episode and the subject does not have to be aware of consistency or historical performance in any single episode. In addition, S wouldn't be able to tell whether the abnormality came from the world or came from herself unless S exchange beliefs within the community. If we reflect on ourselves, we cannot determine whether we are normal observers with a normal set of perceptual abilities without the judgements from other community members. We are confident that how things look, smell, sound, taste, and feel represents their realities because our perceptual beliefs converge with those of other people. It is thus a community's long-term effort to mark its members' consistency and evaluate their knowledge status accordingly.

What this example shows is two-fold: appeal only to internal mental states is insufficient to determine whether certain justification is adequate, and what we need is the validation of the justifiers. This brings us to my second argument. Consider our beliefs justified by reading instruments. Suppose we have a thermometer that works accidentally this one time but failed in almost all other cases. Hence someone who uses this thermometer to arrive accidentally at a true belief wouldn't be ascribed knowledge. However, we usually attribute knowledge to another individual with a similar experience if the thermometer function normally. To know about the temperature in an ordinary context, one is only required to read the thermometer responsibly without being required to check its functionality, that is, calibrating or validating it by checking temperature-reading pairs before using the thermometer. In general, this applies to the devices and equipment we use to secure beliefs every day. We use watches and clocks here and there and a lot of them we only encounter once or twice. It is thus too demanding to

validate the functionality each one of them before using. Yet whether they function normally or not determines whether our beliefs formed on their basis count as knowledge. Thus, if all we measure is what we get access to when acquiring the beliefs, our knowledge status cannot be fully evaluated.

This shows that we need to establish the validity of the justification, reasoning, or evidence we appeal to in our knowledge claims. What we take to be good/bad justification does not come out of the blue, instead, it is the impression we shaped from past experience. What we use as evidence today might be unavailable yesterday, and it might be deemed problematic tomorrow. We cannot achieve this without implementing the idea of reliability or its kind. Take foundationalism, the idea that what justifies our beliefs is a set of basis beliefs that does not require further justification. Typically, foundationalists suggest that we can know "it is a tree" based on justification like "it looks like a tree". But the same reasoning pattern does not work for, say, identifying a work of art as an authentic Da Vinci where it takes more than simple visual evidence. Apparently, we feel secure with our justification in the first case but not in the second. But the justification does not speak for itself. One mistake foundationalists made is that they take the connections between justification and beliefs for granted. We feel like some pieces of information "establish the truth" or "perfectly justifies a belief" because the connection between the information and the factiveness has been deemed extraordinarily empirically reliable. But it is still possible to conceptualize a world in which "it looks like X" fails to serve as evidence for the claim "it is X" for a particular kind X. In a future world where we adopt augmented reality to created holograms of trees that are visually hard to distinguish from real trees, "looks like" justification would cede its power to, maybe, justification from

some technological devices. “Looks like tree” justifies “is a tree” not because there is an inherent connection between these two states of affairs, but because the justification works. It works in everyday life. That is, every time we look at something that looks like a tree out in the fields, it is extremely likely that the object is an actual tree. But when holograms strike, this justification no longer functions normally. Plenty of times when you are presented with a tree-like object, the object turns out just to be a vivid hologram. Similarly, the “looks” justification for Da Vinci’s work of art fails because we can anticipate that it performs poorly due to the number of copies of his work. We seem to be able to evaluate the strength of justification only because we have a general idea of how likely some justification would lead to some beliefs being true. But this impression could not exist without the previous or simulated experience of the same kind. Think about this. If someone brings out a blue flashlight and makes liquid look green and claims the liquid is yellow, we have no initial idea how to make an assessment. This is because this belief-forming process is unfamiliar to us, that is, we haven’t yet had enough experience to establish the track record of the performance of this process. From those experiences, we can measure the performance or how well some justification is connected to certain states of affairs. And this is where reliability comes in. It is the fundamental and constant measure of the goodness or badness of certain evidence without which our evidence wouldn’t possess its justificatory power.

Reliability is the idea to remove these concerns, but we wish to implement the idea without downplaying the insights from internalism. Recall that if I want to know about temperature by reading the thermometer, I have to know how to read it. Thus I must have mental states of the readings of the thermometer. On the other hand, the reliability of the thermometer also affects

attributions of knowledge to me. When stakes are low, it is not my duty to check the reliability of that specific thermometer.[8] In other words, I do not have to be concerned about the quality of the thermometer during the process of arriving at this belief (a child might learn only how to read thermometers but have no idea on the scientific and practical reality of thermometers).

To encompass the internal and external aspects that are equally necessary to my knowledge status, following Nozick, we can model this whole process or think of the whole process of forming the belief as a reliable method.[9] In "Knowledge and the State of Nature", Edward Craig wrote:

Now for the vast majority of human beliefs, if not indeed for all, there is such a thing as the method of acquisition. We do not just have beliefs, we come by them in specific ways... the vast majority of human beliefs are acquired and something, at least, can be said in description of the process by which their subject acquired them. (Craig 1990, 30)

Pay attention to the beliefs we've been acquiring. Not only do we do certain things to arrive at a belief, our actions follow a pattern and we typically have specific instructions for how to arrive at certain beliefs. The way we use watches, thermometers, and fingerprint matching systems to secure beliefs are some perfect examples. We give other people instructions on how to use those things to acquire good beliefs (e.g. This is the way to read a watch. Put your finger

[8] Things might be different in high stake situations. Consider the process of lauching a spaceship where we have to check the functionality of every components in advance.
[9] Nozick is one of the earliest philosopher who adopted the idea of"reliable method" . See *Philosophical Explanation*, pp. 172.

on the fingerprint pad this way. When the computer reacts in certain ways, it means certain proposition is believable). To consider knowledge as a product of reliable methods is to ask us to be explicit about the procedure through which we get to our entitled beliefs.[10] (As to how specific the method is or how exactly should we pick out the method being used in certain episodes, see Question 4 of Section 3.)

We can think of the subject (I) as a user of a method whose only job is to follow certain instructions of a reliable method, whereas the knowledge attributors have to identify what method the subject used and evaluate the method. In this way, we can understand justification as an indicator or illustrator of the method used. Citing evidence surfaces the reliable method from deep water when needed. For example, if I ask someone how he came to believe that it was raining this morning, I would be satisfied with answers like "I saw it" since it indicates that his belief was caused by his direct observation, which is a method deemed reliable. But if we were both standing in the rain and I asked him for justification of "it is raining", he would be confused and had no idea what to say, as the context should have already illuminated the method and so there would be no meaningful verbal contribution.[11]

Again, this reliable method model of knowledge shouldn't disregard mental states as an important component but fundamentally account for it. If we think about the way we acquire knowledge as following a set of instructions, what usually happens is that we have to collect

[10] Note that not all processes of forming a belief feels like using a method. For instance, we wouldn't describe the process of recognizing people as involving some sort of method. But still, we can form a set of instructions and describe the process approximately as a unique set of experience we would like to have in order to form a belief. We use "method" to as a model or an approximation to understand the different set of actions we are taking.

[11] To make it a little bit wilder, maybe the reason we have the convention to ask for justification during a game of knowledge is because we don't get access to other people's minds. If the teacher is able to duplicate the mental activity the student had when she attempted to make an observational claim, it is probably not necessary for the teacher to ask for more explanation to attribute knowledge.

some information before we move to the next step, and all the information along the way should help us get to the conclusion. One easy way to think about this is through mathematical computing. We would like to think of knowledge of math as a product of deductive reasoning, but really we are applying distinctive methods when approaching problems. When calculating the area of a triangle, we are typically choosing between several methods. If we are using the base and height formula, then we must have the correct measure of them in our minds. In other words, we must be able to cite the base and the height when put under challenge. The mathematical method, which is reliable, would then output a belief that amounts to knowledge based on the given information. If we adopt Heron's formula, then we must first record the lengths of the three sides of the triangle. In order to use different reliable methods, there are different sets of mental episodes we must get access to. The failure of accessing those mental states entails the failure of executing certain methods. In this way, we are able to save all the insights we gained from philosophical discussions of justification.

But there is an issue that can potentially interrupt knowledge attribution even we have a reliable method in hand. Although it is not required for the cognizer to be aware of the method used when she acquired the belief, she has to demonstrate that she is familiar with how the method works and its general performance if being asked. In other words, the cognizer must be entitled to use the method; she must establish the legitimacy of her confidence of the method. We cannot be using a random method while it is luck that brings reliability to the method. It is easy to think of counterexamples if this is not a requirement. Consider a reliable computer that discerns whether the individual has a fever. Assume that the computer yields perfectly reliable results. If the computer outputs "1", it means "there is a fever", and the output "0" means "there

is no fever". Now suppose a child C observed "1" from the computer but knew nothing essentially about this equipment nor the technology related. But C constantly observed sad faces from the patients when the computer says "1", and faces with joy when the computer says "0". Further suppose that the only thing C knows about the hospital is that it is a place for diagnosing fever. Thus C inferred that "1" means that the patient has a fever. Hence C had a true belief and formed the belief based on a reliable method. But we still don't feel like ascribing knowledge to C. Let's imagine another doctor S who we would ascribe knowlegde (fever beliefs) to. It is possible that C and S were going through the same mental states when they attempted to form beliefs about whether the individual had a fever. S needn't be thinking of the exact details of the method because she was already conditioned to the correlation between the computer signal "1" and what it represents. The difference between C and S has to do with what they had previously experienced to perform the set of actions they performed and make the set of inferences they made. We would say that S's actions and inferences are well-established by her previous experience, and if S is being questioned, she would be able to defend her entitlement towards the method. But C cannot meet the same requirement. C didn't have the necessary experience with the method nor would she be able to satisfactorily defend her entitlement.[12] Our entitlement towards the method might be problematic also due to the shift of contexts. Suppose scientists established a method M that reliably secures us certain beliefs on earth but had not yet figured out the reliability of M in space. It would then be

[12] One can defend this entitlement not only by reasoning but also by showing. If someone challenges your basic perceptual beliefs, further and further justification just doesn't seem to be helpful. What we typically and reasonably do is to raise plenty of similar beliefs and confirm our beliefs with that of others. Other how we know is unexplainable (see footnote 5). In these cases, we defend our entitlements almost solely by repeating our processes of belief again and again. Demonstrating entitlements by reasoning occurs more frequently in advanced knowledge where the methods are sophisticated (e.g. knowing something by doing a chemistry experiment).

irresponsible to use the method in space regardless of whether the method is actually reliable there or not. Hence people who used M in space wouldn't be attributed with knowledge. In other words, even if M might be objectively reliable, its reliability must be accepted (or understood) by the community in order to have epistemically legitimate confidence of the performance of M.[13]

To sum up, in order to use a reliable method, the cognizer must gain entitlements for using a method. This tells the knowledge evaluator that the use of a certain method is legitimate, that there are grounds for the cognizer to take these actions and make these inferences to form the belief. Recall that children whose observational reports can count as knowledge must be able to be responsible for what they are saying. To achieve this, they must have a general idea of how their observation reports would perform in the community. Imagine that for a child C, we always deny what C purports to be true when C uses her eyes to observe color. Then it would be abnormal if C still confidently commits herself to new color beliefs. In this case, C is not entitled to use her visual perception to commit to beliefs, since C does not have good reasons to believe the soundness of her visual abilities. Only if C got approved frequently when reporting colors can she have grounds to put forward new color beliefs and hold herself responsible for those claims. In our ordinary life, it seems a bit rare for knowledge attributors to challenge the cognizers' entitlements for methods. Indeed, we have been utilizing our perceptual abilities long enough to gain legitimate confidence about them, and if someone is

[13] As to whether the cognizers are responsible for the acquaintance of the contexts in which certain methods are usable, it is a complicated issue, and it has to do also with the stakes. We obviously don't remember all the conditions under which watches would become unreliable (consider variables such as pressure, heat, etc). And it does not really matter in our everyday life. However, for example, if we are in space, we do become more cautious to the functionality of watches and would collect more evidence to approve the use of them.

able to read a watch, it is too unlikely that his confidence of his watch reading came from some utterly non-legitimate source. But if we think about more complicated methods, these challenges come up pretty quickly. If a student were to use an advanced method that's beyond the scope of knowledge in her middle school class, it is reasonable for others to question whether the she does know how that method works. It wouldn't be okay if she merely copied the flow of the method without the ability to understand it. Thus even though she appeared to have all the right steps, her belief still does not amount to knowledge. This "method entitlement" criterion also serves to ground another objection to Goldman's original account of epsitemic reliabilism and all variations of reliabilism that does not invoke the idea of justification.

But access to the justification of the method need not be overly demanding. It's worth mentioning that different levels of effort are required for cognizers under different situations. For a method developed by a community, we have to separate the pioneers from the followers. The pioneers are the early adopters of the new methods, and the followers are the vast majority of the community who employed the methods. To gain entitlements of a novel method, not only do the pioneers have to understand how the method works inside out, they also need to provide legitimate reasons and records that justifies the reliability of the method. Engineers who originally developed clocks would have to do extensive tests to make sure that clocks are the reliable indicators of time. On the other hand, ordinary people, the followers, only need to be informed by these pioneers either directly or indirectly and have sparse and personal experience of the method's performance to obtain entitlements.

Section 2: Measuring Reliability

If we acquire knowledge through reliable methods, then how are methods deemed reliable? What are the criteria of something being a reliable method? What does being reliable mean? In Goldman's terms, being reliable is to produce a sufficient ratio of true and false beliefs or to produce truth X% of the time where X is some number close to 100. (Goldman 1979, 96-97) In Brandom's terms, being reliable is to have "reliable differential responsive dispositions". (Brandom 1995, 900) So the blurry picture we have right now looks something like this: There is a set of situations (call this the reference class). We use the belief-forming method M in these situations to collect a list of beliefs and we compare them to reality. We then compute B(P)/P as the ratio of the beliefs being true using this method. The problem is, how should we constitute the reference class? If P = "This is green", what should be in our reference class? Only green things, or should we add in all sorts of colors? Following Nozick's tracking account of knowledge, I will outline an ideal way to construct the reference class. Let's call the process of testing whether a method M is reliable using this reference class as a T/F Ratio Test.

In "Philosophical Explanations", besides the truth and belief condition, Nozick brings out his tracking account for knowledge:

(1) If p is in fact true, S would believe that p. (positive condition; consistency)

(2) If p is in fact false, S would not believe that p. (negative condition; sensitivity) (Nozick 1981, 172-176)

He then found a problem with this analysis. Since there are many ways we can arrive at a

belief, how we do it has a definitive impact on our evaluation of the tracking conditions. If I hold "the sky is blue" by guessing, then the conditions would not hold. But if I am observing, these conditions would hold. (In case we need a counter example, see Nozick's grandmother example.[14]) Thus Nozick offers a better analysis:

(1) *If p is true and S were to use M to arrive at whether (or not) p, then S would believe, via M, that p.*

(2) *If p is not true and S were to use M to arrive at whether (or not) p, then S wouldn't believe, via M, that p.* *(Nozick 1981, 172-176)*

I believe that this is on the right track. If we ponder what Brandom's "respond differentially to different stimuli" means, it becomes clear that the phrase contains two components. The more evident one is that the subject must separate one type of stimuli from other incompatible ones. Brandom's examplar parrot must be able to respond "this is green" only when things presented to the parrot is green and respond something else otherwise. The less evident component is that the subject must be able to consistently perform the same response in the presence of the same stimuli. That is, when the object is green, the parrot should only respond with "this is green". Human cognizers must be able to achieve these as well (plus all the rest). To recognize a face is to separate this single face

[14] A grandmother would believe that her grandson is healthy by seeing him when he visits her. However, if he was not to visit her due to sickness, she would nevertheless be told that he was healthy by his friends, in order to dispel her distress. We want to say that if the grandmother sees her grandson, her belief can count as knowledge. But if she relies on other people's testimony, she does not know. Thus even though she formed the same belief, the method she uses determines her knowledge status.

from other faces as well as to consistently count this face as the same face.

Nozick's initial account seems to connect the subject S with the tracking condition. But the refined account made it clear that it is essentially the tracking of the method M that allows S to track. It is the uses of different method that result in the difference of knowledge status in the previous example. Focusing on methods is more important than focusing on subjects because methods are at disposal to any member (a standard cognizer) of the community (that are able to use them -- for example, blind people do not have the ability to use a visual process but visually normal people do). The tracking of some method M would then secure knowledge for any number of individuals who hold beliefs on the basis of M. Also, recall that the development of certain methods allows all eligible individuals to get access to some ways of knowing. (e.g. the use of thermometers) The heavier burden is on M rather than on any individual.

To retain the insight of "producing more true beliefs than false ones", we have to make some further specification. What allows us to ascribe knowledge in normal lives is our normative forgiveness for mistakes. If I exert my memory and made a mistake once, then there is salient room for mistakes when it comes to future beliefs based on memory retrieval. Thus, strictly speaking, the tracking condition is not satisfied for beliefs like this. Nevertheless, this shouldn't ruin my opportunity to have my beliefs based on memory counted as knowledge. Under normal contexts, the community would not immediately distrust an individual when she reports incorrectly just once. Sometimes it just so happened that we made a little mistake. When we are children, our parents and teachers are our sources of knowledge. Even though they are not always right, we can still rely on their

testimonies as long as they are right satisfactorily frequent. (The method children use is basically to believe what the teachers and parents assert to be true.) This simply means that we need to allow some degrees of freedom in the performance of our methods.[15] So rather than "S would/ wouldn't believe that P", we should say "it is very likely that S would/wouldn't believe that P".

So how are these two conditions necessary? Why would violating any of the two conditions rule out attributing knowledge? Let's dive into the positive condition first. Here is one example to lead the way. We rely heavily on calculators to perform arithmetic. In this case, I typed in 7^5 and the output is 16807. Then, normally, people would say I know 7^5=16807. The positive condition must be met because every time I activate that method, that is, I perform the input, the output from the calculator must be 16807. If not, then there is probably something wrong with the calculator and I got the right result 16807 out of luck. For example, if the calculator is a random number generator and it just so happens that the first output is 16807, I do not know 75=16807. Similarly, when the temperature is 32 °F, my thermometer should present a reading of something around 32°F every time. The positive condition serves to secure the consistency of a method's performance. If I am caused to hold a true belief by a method but the method is inconsistent in the results it produces, namely the method may as well offer me a false belief easily, then my belief does not amount to knowledge. But, as I argued in the last paragraph, we don't need perfect consistency -- a minor inconsistency is tolerable.[16] Human cognitive systems, in general,

[15] For other examples, consider scientific claims (that we take to be knowledge) based on hypothesis testing.

[16] Even though, for artificial devices like a thermometer or a calculator, a single mistake usually

are indeed relatively unstable. Perception deceives us sometimes, but we are okay with it as long as we enjoy consistency under most circumstances. However, for a subject S, as inconsistency increases, we are less and less willing to attribute knowledge to S. Consider the eyesight test. If the examiner pointed at a single letter multiple times but I only get it right once or twice, then my consistency is fairly low and I wouldn't be attributed with knowledge. But if I recognized it mistakenly just one time out of many, it didn't really matter. Thus knowledge attribution can only take place when the track record of the cognizer's beliefs using that method meets the standard.

DeRose has spread doubts about this True/True subjunctive. (DeRose 2004, 29-31) We think that knowing P requires P to be true. So what does it mean to suppose that P is true when P is already true?[17] The answer is we are considering other similar situations, either historical records or the product of counterfactual thinking, where P is true. For example, when we first evaluated the reliability of any computational machine (or basically any initial evaluation of equipment), we generated a track record to validate the consistency of the outputs. (e.g. we tested that when the desired results are all 1's, the computer has to produce 1's all the time.) When we began to attribute perceptual knowledge to our children, we can do that only because they previously did well, which means they have been consistent in their performances. Other times, however, our evaluation leans towards the hypothetical or imaginary side. In the calculator case, I would not, apparently, check the

translates to unreliability. But for other techniques such as DNA testing and radiocarbon dating, we would always have a confidence interval associated with a conclusion, meaning that there is a slight possibility that the conclusion is incorrect.

[17] The consistency condition is: If P is true, S would believe in P. But "P is true" is already a condition for knowledge. It thus seems that the subjuntive "if P is true" is misplaced.

answer 16807 multiple times to verify consistency. But if I am to perform the same task multiple times, it can be anticipated that the outputs are all 16807 since these computations can be expected to be deterministic. The indispensability of this positive condition can also be seen in science. If a scientific claim is the product of a set of experiments, and if scientists were to believe it, they should've believed the statement everytime they conduct the same set of experiment. The violation of consistency directly poses a threat to the knowledge status of certain beliefs. An established psychological claim can be called into question if the data of a new proper experiment under the same procedure rejects the same claim. Again, this measure of consistency ensures that it is no accident that we arrive at this belief in this single episode.

For sensitivity or the negative condition, notice that the arithmetic examples would make the condition redundant, as there are no cases in which this mathematical proposition is false. But things are different for contingent truths. If I simply claim that everything is green, it allows me to be correct whenever the object is actually green. But I wouldn't correctly be ascribed knowledge since in those cases where things are not green, I still believe that they are green. The following table helps illustrate the need for sensitivity:

Suppose on the fruit market of our community, there are only three kinds of fruits: apples, bananas, and oranges. For each kind of fruit, there are 1000 of them. Let our subject S identify each one of them.

Table 1

		Reality		
		apple	banana	orange

Belief	apple	999	0	1
	banana	0	1000	999
	orange	1	0	0

If you look at the table, even though S does very well when being presented with bananas, we wouldn't say that S can identify bananas because S also takes oranges to be bananas. But when it comes to apples, it is safe to say that S knows what apples are.

There is a competing Safety account provided by Ernest Sosa which seems to be the contraposition of sensitivity:

If B(P), then P would be true. (B(P) stands for believing that P) (Sosa 1999, 142)

We have made it clear that there must be room for mistakes for the simple fact that none of our actual ways of knowing can nail the truth 100% of the time. One failure of applying my memory capacity would not ruin the credibility of my memory. Then the safety account must be interpreted as:

If B(P), then it would be very likely that P.

whereas the sensitivity account is:

If ~P, then it would be very likely that ~B(p).

This interpretation made it clear that sensitivity and safety are not really contrapositives to each other. (The contrapositive of sensitivity is “If it is not likely that ~B(p), then P”.) But if the safety analysis must be put this way, then it would permit some counterexamples when the actual frequency of P and ~P are vastly uneven. The lottery case is a good example. If we believe that we would not win the lottery, it is very likely that our belief turns out to be true. Thus safety is satisfied. But our intuition is that we don’t know the outcome. Here is a good illustration using our fruit model. Suppose we are in a world where the number of apples is significantly larger than the number of bananas. Consider the following table. This time S has to identify 10000 apples, 10 bananas, and 1000 oranges.

Table 2

		Reality		
		Apple	Banana	orange
belief	apple	9999	10	1
	banana	0	0	0
	orange	1	0	999

The safety account starts out by collecting all the cases in which the subject S believes “this is an apple”. Then the table would suggest that it is very likely that the object is indeed an apple, since there are way more instances of apples than those of bananas. However, notice that S takes all 10 of the bananas to be apples, which indicates that S can’t really tell bananas from apples. Since S does not rule out the possibility that the object is a banana, S doesn’t really know that the object is an apple. The advantage the sensitivity analysis provides is that we force

different tests for situations where ~P is true and consider their performances separately. We don't think we know the lottery outcome because when we consider the salient possibility that we might win, our belief-forming method (always claiming we lose) performs poorly. In the table above, S is insensitive to banana when it comes to identifying apples, and this is exactly why we don't feel like attributing knowledge to S. The safety account does not itself provide a mechanism for forcing the even frequency for every alternative. (That is, for example, force the reference class to be something like 1000 apples, 1000 bananas, and 1000 oranges.) Besides, when the frequencies are actually even like what happens in the first table, the safety analysis does not seem to have an edge over the sensitivity analysis.

What induced the safety account in the first place might be the tension between the sensitivity analysis and epistemc closure. Indeed, a primary concern with the sensitivity accout is the simultaneous violation of epistemic closure, also known as closure under known entailment: But in the following discussion, I argue that they are not really incompatible. Here is the Closure Principle, also known as the closure under known entailments.

(Closure) If S knows that P, and S knows that P entails Q, then S knows that Q.

This principle should be intuitively sound. For example, I know that the leader of the Nazi Party is Hitler. And if I think about what that implies, it is easy for me to conclude that no one else was the leader of the Nazi Party. But the sensitivity analysis seem to be in conflict with this principle. Briefly, if closure is satisfied, then the sensitivity requirement isn't and vice versa. For example, under closure, I must be knowing that this thing in front of me is not a Martian

machine that looks just like an apple if I claim to know that this is an apple (by looking). However, my belief that this thing is not a Martian machine that looks like an apple is insensitive (by looking), which indicates the failure of tracking. Nozick is well aware of this situation and so denies closure. Nozick brought up the following argument: if S knows that S is in Jerusalem, it is not required that S knows that S is not "floating in a tank on Alpha Centauri being stimulated by electrochemical means to believe that 'S is in Jerusalem'". (Nozick 1981, 207)[18] For Nozick, we don't know the improbable possibilties. But this situation does not have any bearing on whether our normal empirical beliefs can be attributed to us as knowledge. Nozick believes that sensitivity is epistemically more important than closure and thus does not have an issue with the failure of closure. On top of that, Nozick argues that sensitivity analysis does not have to include fantasies like this. It only needs to work within the closest possible worlds around the belief P. In other words, since fantasies are in worlds that are remote to P, it does not have to be considered.

I agree with Nozick's "closest worlds" analysis, but not his take on closure. This is my version of the closest world analysis, which can maintain the closure principle:

Suppose P_1, P_2, P_3,... are mutual alternatives such that P_n entails $\sim P_1$, $\sim P_2$,..., $\sim P_{n-1}$, $\sim P_{n+1}$,... Let S hold the true belief P_1 in a context C. It can be the case that the belief-forming method M adopted by S only rules out P2 through Pn, given that all the other alternatives are excluded by their unlikeliness in the context C. Thus we can appropriately attribute ~P2, ~P3,..,~Pn to

[18] Note that "S is in Jerusalem" is incompatible with "S is floating in a tank on Alpha Centauri" and so they are alternatives to each other. For the Closure Principle to be satisfied, if S knows the former, and if S knows that the former entails the latter (which is easily knowable), then S must've known the latter.

S by the sensitivity of the method M, and attribuite ~Pn+1, ~Pn+2,... to S by their contextual improbability.

To sum up, my argument is that the method M only needs to exclude P2 through Pn and all the rest are irrelevant possibilities that are deemed unlikely in the context C. Thus S can conclude all propositions entailed by P and closure is preserved.[19] Note that the Closure Principle is about the subject S, whereas the sensitivity analysis should be about the method M. In preserving closure, our method M only takes a part of the burden.

Why? The idea of "closest worlds" is not a philosophical compromise but, as a matter of sociological fact, how humans get things operate. Belief-forming methods developed and applied by communities are, more or less, heuristics that do not cover or mean to cover all logically possible situations. That is, our methods are only sensitive within a certain range of alternatives. The sensitivity analysis should be applied in a local context. We can know the speed of a car by looking at the gauge reading (the method is reporting whatever the gauge reading is). If the gauge says 60 mi/hr, I believe it and I would normally be attributed knowledge. Nevertheless, we are not sure whether traveling at the speed of light would produce a reading of 60 mi/hr, as we have never done the experiment with that gauge nor do we have any idea what could possibly be anticipated in terms of the gauge when we are in that situation. However, it is plausible for us just to ignore that possibility, even though the belief-forming method does not preclude that from happening, for it is not the responsibility of this specific method to rule out this possibility. The unlikeliness of this situation has already been taken care

[19] There is no contentious issue if Q is not the negation of an alternative of P.

of, that is, in our current context, we shouldn't be taking traveling at the speed of light as a serious alternative (based on induction). In this case, the gauge has to be sensitive only within the normal range of possible vehicle speed.

So what about the other part of the burden in preserving closure? The answer is we reject certain possibilities because it is appropriate to suggest their negation based on being reasonable, not because of the method we employed. We can apply this idea to a famous closure denial example from Dretske: We look at the sign which says "zebra" in front of the pen in the zoo. We look up and see a zebra-like creature in the pen. Dretske argues that the Closure Principle does not hold because even though one can know the creature standing in the pen is a zebra, one does not know that there is not a cleverly disguised mule (CDM) in the pen. (Dretske 1970, 1016) Dretske believes that we don't need to take care of this possibility since it is irrelevant to the current context. I have no problem with this. So what we really disagree on is the intuition of whether we know the absence of CDM. We do have evidence that this is not happening. But this evidence is different from the sign or our visual input because (1) it is not collected during the scenario (2) we don't have to be thinking about the evidence during our inquiry. Based on previous exposures to zoos, we should assume that the zoos have been running normally and have been honest and so it is unlikely that we can be tricked when identifying an animal in a zoo by normal means. Thus fake identification does not even have to be a concern in our inquiry. (If, however, it has been disclosed that some zoos actually implement fake substitutes, then we might need to consider taking our assumptions back. But this would involve a contextual shift.) The context of our inquiry enables us to assert the negation of these improbable possibilities and consider only a restricted pool of alternatives.

We thus require our belief-forming method to consider only these cases. (e.g. The method needs to preclude the object from being other ordinary animals in a zoo.) And if the method is locally sensitive, the Closure Principle is satisfied.

We might be hesitant to say that we know the negation of irrelevant possibilities. From the perspective of language games, the whole issue of whether we can say we know them really comes down to this: our normal inquiry should not be interrupted no matter how we choose to word our sentences. We don't allow, for example, the logical possibility that there are billions of humans living in the core of the earth to interrupt our conclusion that there are approximately eight billion people on earth. We are entitled to dismiss them as irrelevant or assume their falsity if we intended to join in this language game, otherwise, our language games would face a complete break down. So the tension does not seem to be that important. The following two takes on closure should both be acceptable:

A: We assume that irrelevant possibilities are false and their negations are true. We just can't say that we know them. Thus closure does not hold. But we might have a Restricted Closure Principle that holds: If S knows that P, and S knows that P entails Q where Q is in the close possible worlds relative to the context C, then S knows that Q.

B: Based on experience, since there is no evidence that the irrelevant possibilities can occur, we are in a good epistemic position to appropriately say that we know that they are false. Thus the Closure Principle holds.

For me, I am inclined to take B over A because in language practices, I see no problems with people committing to the negation of irrelevant possibilities based on their common sense. For example, it is logically impossible to establish the non-existence of something and so some of the negation of these propositions belong to the irrelevant possibility family. However, it is appropriate for people to reject some of them. (e.g. There is no such thing as a unicorn.) For every case of murder, it remains the logical possibility that it was done by some extraterrestrial forces. But no judicial procedures could be achievable and no reasonable conclusion could be made if we wished to collect conclusive evidence against these possibilities.[20] Besides, it seems very acceptable, in those contexts, to say something like "It is impossible that aliens committed this murder." So much for the tension between sensitivity and the Closure Principle.

Now that I have defended a version of Nozick's sensitivity account, there is one final concern, and that has to do with how to measure the T/F Ratio under the sensitivity analysis. I argue that, for all relevant alternatives, we need to consider every single alternative separately. This should make sense because, for instance, suppose I know this is John and John has a twin Jacob, I can obviously be very good at differentiating between John and every non-Jacob-non-John, but if I can't differentiate between John and Jacob (who is in the context), then I don't know for sure this is John. Thus even if I have a good performance record when the measurement combines all alternatives together, it is still not enough for me to know this is John. We can also refer back to Table 2. If we combine the banana and orange cases together, the sensitivity analysis would yield a high ratio (98.9%). But clearly we wouldn't say that the

[20] The term "beyond a reasonable doubt" is used as a standard of proof required to convict a criminal in legal systems. We can think of all the unreasonable doubt as the set of alternatives that can be reasonably ignored in order to reach a reasonable conclusion. This should allude to our language in general. If no reasonable assumptions have been made, no further moves can we take.

subject S in that scenario satisfy the sensitivity condition. If we hold separate sensitivity tests for banana and oranges, the problem disappears.

All these considerations discussed in this section lead me to propose a version of reliabilism that saves a version of closure, takes epistemic context into account and that rests on what I call a T/F Ratio Test. Here is the final statement of the test:

The reference class of the T/F ratio test of a method looks like this: A good number of P_1, P_2, P_3,..., P_n *all of which constitute situations that we normally encountered in our context such that* P_i, P_j *are alternatives to each other for all i,j. All* P_{n+1}, P_{n+2}, ... *are irrelevant possibilities and do not belong to the reference class. Let* $N(B(P_n))$ *be the number of instances of believing or affirming that* P_n, $N(\sim B(P_n))$ *be the number of instances of not believing or denying* P_n, *and* $N(P_j)$ *be the number of instances of* P_j *being true. Assume that* P_x *is true. A method M for arriving at* P_x *would pass the T/F Ratio Test if and only if* $N(B(P_x))/N(P_x)$, $N(\sim B(P_x))/N(P_1)$, ... $N(\sim B(P_x))/N(P_n)$ *are all satisfactorily high.*[21] *(Notice that the first fraction signals the consistency and the rest signals the sensitivity.) For all* P_{n+1}, P_{n+2}... *that are remote in the context, we are entitled to say (or at least assume) their negations are true.*

Section 3: Approved-list Reliabilism

Now that we have a model for the T/F Ratio Test of a method, the natural next step is to suggest that a method can be called reliable iff it passes the T/F Ratio Test, and we can attribute knowledge by considering whether a subject S adopts reliable methods to form a belief. But

[21] The range of these fractions is between 0 and 1. A satisfactorily high ratio would naturally be something close to 1, but it has to do with the context. We will discuss that in Section 3.

this is too quick. Let's consider this objection raised by Laurence BonJour. (1980, 53-73) Think about Henry the clairvoyant who has a perfect record generating beliefs based on his clairvoyance faculty. BonJour suggests that Henry's belief does not count as knowledge (since Henry does not have justification), but reliabilism must say Henry is.

This begins our journey of seriously examining how we use the term "reliable method". Under what circumstances are we allowed to deem a method reliable? The example serves its purpose by stipulating a perfect track record for Henry's clairvoyance, thus it would seem to count as a reliable method. But wait, don't we believe implicitly that clairvoyance does not count as a reliable method? If we firmly believe that clairvoyance is reliable and trustworthy, why are we reluctant to ascribe knowledge to Henry? Think about how this leaves us with a very odd situation. On the one hand, we believe that Henry does not know. On the other hand, for another individual Alice, Henry's beliefs are perfect testimonies. Alice could simply recognize Henry as a reliable blackbox where the outputs are true beliefs. This is no different from a layman's use of a thermometer. Since Henry's beliefs based on the clairvoyance faculty are always true, Alice has good grounds to use Henry's beliefs as a reliable source of information. As a result, Alice can be a called a knower. To push it a little bit further, Henry, based on Alice's testimonies which are based on Henry's testimonies, is now able to be said to know. We might contend that we know how thermometers work but not Henry. But we don't have to. If all experts and literature on thermometers perish, we can still attribute knowledge to people who use the thermometer. Imagine if several laymen evacuate from the earth where everyone else was dead. Even though they do not have access to the detailed information of how the devices they carry work, we wouldn't say that they are immediately deprived of

entitlements to their beliefs based on those devices. We might contend that Henry doesn't have "evidence". But we can make the case that the faculty produces different signals for Henry when it comes to different situations, and Henry is caused to believe different propositions based on different signals. This is similar to someone with an auditory sense far better than everyone else who can be caused to believe the existence of sound when everyone else cannot. While we tend to attribute knowledge to the latter, we are not equally welcoming the former. This is a product of unfamiliarity and distrust, and probably, tradition. Currently, we don't believe in clairvoyance; it does not have a good reputation; we don't see a reason to believe in it; we don't know how it operates and so on. If, on the other hand, clairvoyance has been real and a good amount of research has been done to explain how it functions and it has been delineated as a superpower in a documentary that folks love, our intuition and thus knowledge evaluation would surely change.

This points to something important. There is a gap between being a method that has objective reliability and asserting that a method is reliable. Since we are focusing on language practices, we want to make sure our account depicts only what is accessible and accepted for a community so that language moves are plausible. An objective reliable method, if any, can fail to be recognized by a community and thus has no impact on language games attributing knowledge. The use of such a method would be marked as irresponsible.[22] We can think of a genius who developed a method for determining whether there are living organisms in a given region and the method might be able to generate true beliefs sufficiently frequently. But if we cannot fully understand the genius's program and there are no other ways in which we can

[22] We can also say that the confidence of using this method comes illegitimately. See Section 1.

validate that there are lives in a given region or not, we cannot trust that method regardless of that method's true performances. We are then reluctant to mark beliefs formed by that method as knowledge, namely, the method cannot play a role in language games. We shall observe in this case that for a community to attribute knowledge to the subject, a community must itself be able to understand and evaluate the method for knowing that allows him to make that claim. What is going wrong in the clairvoyance example is that it might be the case that clairvoyance is objectively reliable, but that does not imply it is immediately appropriate for a community to assert that clairvoyance is a reliable method.

Our task is to fill in the gap between passing the T/F Ratio Test and being a reliable method that takes up a position in language games. Elizabeth Fricker proposed Approved-list Reliabilism as a novel way to think about reliability-knowledge connection due to her concern that actual or objective reliability is over-demanding. (Fricker 2016, 109-112) For Fricker, an approved list of reliable methods is a list of methods developed by folks throughout their empirical life such that those methods are deemed to be reliable upon assessments and calibration. We then ascribe S knowledge based on whether S applied methods from the approved list to arrive at P such that P is in the range of what the method is able to establish. We can say that the perception and memory of ordinary people are generally on the list, while wishful thinking and guessing is absent from the list. Fricker suggests that this is only a tentative account that requires further illustration. I argue that this is a promising way to formulate a reliability theory because of the advantage it provides in explaining how we talk about knowledge. We need to answer these four questions:

1. What does it mean to have a method approved?

2. How do contexts and stakes play a role in approved-list reliabilism?

3. What are the criteria for approving a method?

4. In cases of knowledge attribution, how should we identify the methods cognizers adopted?[23]

Question 1: What does it mean to have a method approved?

By approving a method M, we are committing ourselves to the reliability of M in a context C in generating a set of propositions S for the eligible cognizers. Then we have the privilege to attribute knowledge by (primarily) examining whether an approved method has been adopted by those cognizers. However, M itself, the context C, and the set of propositions that M can be responsible for is constantly being evaluated in light of new information. If it is disclosed that there are certain problems with M, C, or S, M might be suspended from being approved or certain modifications has to be made to M, C, and S. Approving a method M is undertaking the commitment that for a set of propositions S, M passes the Ratio Test for every proposition in the set. This seems a lot. But for the community to make a commitment is, again, a reasonable provisional conclusion. We don't have to conduct rigorous T/F Ratio Test for every proposition in the set, for the test is just an ideal and paradigmatic model. Let's use the speed gauge for example. The set of propositions S that can be secured by reading the gauge (only) is S={*n* | The speed of the vehicle is *n* mi/hr}. We can't be expected to test all of the propositions in S.

[23] This question is not directly related to the approved-list reliabilism. However, this novel account is able to offer us a new perspective on this issue. Specifically, it offers a reply to the Generality Problem proposed primarily by Earl Conee and Richard Feldman.

Rather, we usually pick $\{n_1, n_2, n_3, n_4...,n_j\}$ as our facts and see if our beliefs based on the gauge align with them. Then we reasonably assume that every one of ni's would be passing the test. Notice that if we find out that some n_i failed the test, then our assumption should be revised and n_i would be removed from the set S (and possibly other propositions such that the situations are more demanding than n_i) and the method would be restricted appropriately. For instance, we might find out that our gauge does not perform well when speed exceeds 200mi/hr. The set would then shrink if that happens. Sometimes such failure might lead to the reevaluation of the method as a whole. Finally, we are also committed to the general context where a method can be used. We might find that our method does not work in extreme environments such as strange magnetic fields. Visual perception does not work very well when the individual is drunk. We might also improve our method so that we can fit the method into more situations. For cases where a method faces severe challenges such that it might lose its position, think about our beliefs generated from watching video recordings in a context where deepfake is popular and not restricted. We can also think of how people's view on the trustworthiness of the news media changes.

Approved-method provides us with the criteria for the legitimacy of knowledge attributions assuming that there is no additional information which indicates the belief being false, or the method generating the belief is problematic in this context.[24] This structure of approved-list reliabilism can be illustrated by considering how it explains our intuition about Gettier cases. In Gettier cases, we hold accidentally true beliefs, but what's more important is

[24] This thought is influenced by Robert Brandom and Michael Williams' "default and challenge" model of justification. (Williams 2001, 149) On that model, a subject is entitled to hold a given (non-inferential perceptual) belief as long as there are no reasonable or justified challenges to it.

that we think we secure those beliefs by non-accidental means. Thus what we need to explain is not only that we cannot be attribute knowledge when Gettier strikes, but also our immediate perception that we are in a strong position to attribute knowledge. Let's say that S believes that 'the ball game starts at 5 pm' by checking the official website. However, the ball game is designated to start at 4 pm and so the website made a mistake. But the game is somehow postponed to 5 pm so S accidentally gets it right. Why do we think that S is justified or is in a strong position to know? Because S is relying on a credited website, which indicates the occurrence of an approved method. What this means is that if all that we know is the first half of this Gettier story, it is definitely appropriate for us (attributors) to attribute knowledge to S, as this is the privilege of having a method approved. But this privilege can be taken back in light of new information. When attributors hear the second half of the story, it is then salient that the method is unreliable with regards to S's belief. In particular, since it is an accident that the game starts at 5, the method would not pass the sensitivity test. Thus we would take back our knowledge attribution. However, this case itself does not immediately pose a threat to the approved method overall. We can still trust the website by ignoring this unusual case. But if this kind of case (they can be Gettier, or they can just be normal false beliefs) begin to pile up, then it's more likely for us to mark the website as discredited and suspend its privilege and thus block beliefs generated by that source from being knowledge.

We already see how it deals with the clairvoyance case and any case that brings up a weird reliable method that our community does not know how to handle.

Fricker believes that this approach could remedy her concern that objective reliability is

over-demanding.[25] I believe this is a good point and it points towards something important. Whenever we discover or develop a way of knowing, we cannot be expected to perfectly secure the reliability of the method. Logically, there always remains the possibility that there is a sequence of test cases that can potentially render the performance of the Ratio Test unacceptable. A person who suffered from red-green color blindness can be lucky enough to pass the color vision test at school and be deemed to be a knower with respect to relevant propositions, even though it is very likely for the defect to be revealed. The reality for approving a method is that we only have a limited amount of resources to draw a conclusion about our methods. Even though we might be making objective mistakes, we are still being appropriate in making the decisions. Here is a fascinating example. Fermat's Last Theorem has now been proven. But the first "proof" has a problem. Andrew Wiles submitted a proof of Fermat's Last Theorem in 1993 and it was found in the peer review that the proof contained an error. Suppose nobody ever noticed that error and the proof was vetted and published like the revised version did in 1995. Then we would have entitlements for statements like "Wiles knows Fermat's conjecture is true", "We know Wiles proved Fermat's conjecture" and so on, even though, in reality, those claims are not generated by objective reliable methods. However, our social practices force it that once the proof is published and unless somebody made a valid argument against the validity of the proof, we continue to hold our beliefs towards the proof. If instead, we demand objective reliability, then no vetting processes would ever be adequate to establish that objectivity and so we are paralyzed and cannot be ascribing knowledge of the same kind. The fact that we terminate the investigation and started ascribing knowledge once

[25] For some examples from Fricker, see Fricker 2016, 107-108.

our effort reaches a certain point is the key to how language games are operated. The objective reliabilism is too demanding in the sense that if we are not allowed to ascribe knowledge because of potentially undiscovered errors, then we are equally not allowed to ascribe knowledge in similar cases, which contradicts our intuition that we should.

The advantage approved-list reliabilism provides is that when a reliable method is on the list while in fact, it would not be in light of new information (e.g. the Wiles case), we nevertheless appropriately attribute knowledge before we face the challenge. The approval is conditional unless and until contrary evidence emerges. This tackles all challenges that take basically this form: The world is weird in the way that forces no reliable methods. We are nevertheless justified in those beliefs that are formed in unreliable ways. The Wiles case already looms over my reply here. The absence of modal reliability does not imply the absence of approved reliability. Approved reliability commits us to assumptions that have to be made in order to make moves in language games. If we live together with an evil genius, for example, our beliefs about ordinary objects are justified because we trust our standard perceptual processes to be reliable based on previous social experiences. Thus we do have those methods on our approved-list and are thus making knowledge attribution moves like we normally do. Another interesting example. At this moment, we are inclined to believe that we know the battery level of a phone based on the battery icon on the screen. It might just be the case that, unbeknownst to all of us including the engineers that designed and tested this feature, there is a magic power that controls the on and off of all phones. It is a favor the magic power does for us that the phones are turned off once the icon displays digits close to 0, while the digits on the phone do not represent anything but the output of an algorithm

implemented by that magic power that reduces numbers randomly in the stream of time. Since this whimsical scenario is not known by any member of the community, but the community already establish this method for knowing battery levels as a reliable method, the scenario would not affect language moves concerning knowledge attribution unless it is discovered.

Question 2: How do contexts and stakes play a role in approved-list reliabilism?

The approved-list reliabilism also enables us to address the contextual picture of knowledge attribution of knowledge nicely. Some methods can only be legitimately used in certain contexts or situations. We cannot rely on thermometers if they are under abnormal pressure; we cannot trust someone's testimony when she is drunk; and we cannot trust our own eyes in a world with vivid holograms. Besides the fact that methods might fail to produce true results frequently in certain contexts, some methods are inapplicable in certain contexts simply due to disapproval from the knowledge attributors. In middle school, let's say, we are entitled to say "a rational number plus an irrational number equals an irrational number" just by listening to the teacher's words. This method is perfectly reliable since the statement is true and the teacher is seen as some sort of epistemic authority. But if we are in a college level course, we may be asked to justify how we know it without the help of others. Thus relying on testemonial information isn't an applicable method in the college context regardless of whether that method is actually reliable.

Recall that I require the test score of the T/F Ratio Test to be "satisfactorily frequent". We cannot have a precise measure because the standards of approval is different in different

contexts. The bar for what counts as a good performance of a method would generally be raised when stakes get higher. Let's consider what happens on the court versus what happens at school. For the judge and jury to know the occurrence of some events, we might need multiple testimonies and plenty of evidence, thus believing based on the word of one witness is not an approved method. While at school, it is usually good enough just to believe another relevant individual's testimony. Thus when stakes are higher, for the same proposition, we require superior performance (higher score in Ratio Test) from methods and so we can expect the list of approved-methods to shrink.[26]

Stakes are only one way of explaining the demand for better method performances. Different social environments can be another factor. There is no problem for one community of people to be more careful and rigorous than another. There is no problem for some groups of knowledge attributors to be more harsh on the method used than others. Imagine a community that is into collecting instances of optical illusions. They can thus be more sensitive to naked eye observations and be more likely to disapprove these methods. Or they may require a more rigorous T/F Ratio Test with a larger sample size, making it hard for these methods to be approved.

Question 3: What are the criteria for approving a method reliable?

So what, if any, are the criteria for legitimate approval of methods? Recall that the T/F Ratio Test sets standards for what counts as a reliable method. However, a community may never

[26] Stakes might also influence which method the subject chooses to rely on. But the restricted method pool for subjects does not have to be correlated with the restricted method pool for the knowledge attributors. A subject might proofread for more times when the work is extremely important. But this doesn't mean that a knowledge attributor should require the same level of rigor.

get to the objective test score of a method, since no matter how many trials the community has performed, the real test score could be otherwise. What the community does, much like taking a sample from a population, is to perform a provisional or local T/F Ratio test as an estimator of the actual T/F Ratio Test.[27] The provisional T/F Ratio Test must have satisfactorily high test scores, and there cannot be evidence indicating the potential failure of the test. This test is not to be understood as something very different from the test model I introduced in Section 2, it's just that the reference class (or the sample) of the test is limited for realistic reasons and can always be expanded and modified. Let's consider an emperor who brought out a new method for securing beliefs and deemed it reliable. Suppose the emperor was a pretty reliable source before and so was widely trusted by the commoners; the latter have no problem trusting that the method works. At this moment, the provisional tests contain nothing but a reasonable envision of its good performance. When the method starts to be employed, the reference class of the provisional test expands, and depending on how the test score plays out, we can have different evaluations of the methods. If it turns out that the method does not lead to true beliefs sufficiently frequently, then the method should be disapproved.

Other than passing the provisional T/F Ratio Test, I argue that, like other decision-making processes, there are no clear criteria. However, this does not mean that I am getting away with explaining why this is the case or why should we look at it this way. There is no obligation for a community to take a method to be reliable no matter what the objective T/F Ratio Test Score of that method is. Think about if there really is a reliable fortune teller who

[27] Obviously the community wouldn't have any intention or perform any action. But we can think of the community as doing such as such to capture the general picture.

makes specific claims (e.g. Exactly *n* people would say hello to you on the next Saturday) about people's future and get it right every time for 100 times. Unbeknownst to everyone else, she establishes those truths by being informed by God. Even if the method had a perfect record and will potentially continue to do so, people can still be expected to be reluctant in buying this specific fortune-telling given the crooked fame of general fortune-telling methods, and that is totally fine.

It may be hard to accept the idea that we might be indecisive towards our belief-forming methods. However, consider our encounter with novel methods, for example, machine intelligence. We are currently at a stage where algorithms can do a better job than humans at identifying objects, diagnose symptoms, and making classifications and they would only become better. While most of them are linked to decision making and automation, there are cases in which we are knowing based on these algorithms. If I pick up a lost phone and the phone happens to have a facial recognition system (FRS), it can help me to know whether the person who claims to own the phone is the real owner. FRS can connect the individual in front of you to the information stored elsewhere. We can imagine that, in the future, people would have a chance to immediately grasp the information of someone upon having her processed by an FRS. Belief-forming methods relying on FRS would be approved by the community. Of course, it depends on the algorithm being accurate enough. But here is the thing. If an algorithm has the same level of performance as FRS, but instead of predicting whether this is your face, it predicts your political preference, would you agree as easily that the person who runs the algorithm knows your political preference? Thus even though these algorithms are similar

in nature and pass the T/F Ratio Test equally well, the latter is predictably more contentious than the former when it comes to knowledge attribution. The more important point is that using this type of algorithm to generate beliefs involves accepting new methods that generate a lot of suspicion. Even though for example, computers are better at identifying certain objects than some people, we are still more willing to attribute knowledge to those people rather than someone who interprets the output of the computer. It takes time and exposure for the community to get used to this new concept and what it means for us. Think about the following scenarios:

An algorithm A that does the following thing objectively: it takes a picture of the room as an input and then outputs a code that corresponds to an existing thing in the room.

Question: Can we be attributed knowledge about what is in the room based on reading the output of A?

Scenario 1: We have a software engineer explaining to us how A works to identify humans in the picture. He then presents the good performance of the algorithm and demonstrates its capability by letting everyone use it.

Scenario 2: We just recently discovered A written in an indecipherable programming language and it seems that there is a correlation between the output and the things in the

room. Studies have shown that in several trials, whenever there is a person in the room, we receive an “0”, and in the absence of people, there is no “0”.

Scenario 2*: A company bought A and turned it into a product for security purposes. In several years, the products have been extremely accurate in detecting people and it has not been reported once that it made a mistake.

Scenario 1 is the case where knowledge attribution is likely to happen due to exposure to the principle as well as the performance. Scenario 2 is where there could potentially be a divergence. Some people might decide not to trust anything they take as beyond their own reach and so they would not buy anything the computer says. Others might focus solely on efficiency and accuracy and so they are willing to take the computer to be reliable as long as the result it produces is still satisfactory. Scenario 2* is the case where the algorithm interacts with the community for a long time and has built a good reputation. I believe it is more likely for people to approve this method than in Scenario 2. This concludes my argument that there are no clear criteria for when a method is approved.

Question 4: In cases of knowledge attribution, how should we identify the methods?

How can knowledge attributors identify the methods adopted by the cognizer in the first place? How specific should the methods be? It turns out that there is a philosophical discussion centered on these issues. The Generality problem is the problem of specifying what method

has been used in forming a belief. Specifically, it is the worry that the specification of the method could be too general and too narrow. For example, “using perception” and “using visual perception at this distance at this angle at this time in a year” both specify a method that we might be using, one being very broad and the other being very specific. They have different levels of reliability and they cannot both be representing what method we are using. In “The Generality Problem for Reliabilism”, Earl Conee and Richard Feldman highlighted this issue as a primary concern for reliabilism. They provided an example of a subject who looked out of the window one sunny afternoon and saw a maple tree. We think she knows the existence of the maple tree. But what exactly is the method used? For Conee and Feldman, the process of observing a maple tree involves a lot of subprocesses. They wrote:

> “The token sequence in our example of seeing the maple tree is an instance of the following types, among others: visually initiated belief-forming process, process of a retinal image of such-and-such specific characteristics leading to a belief that there is a maple tree nearby, process of relying on a leaf shape to form a tree-classifying judgment... The number of types is unlimited.” (Conee & Feldman, 1998)

There do seem to be a million of ways to characterize what the subject has been going through, general or specific. More importantly, different types of processes would have different measures of reliability. If we just consider general visual belief-forming processes, they seem to be less reliable than processes that use a maple-leafish visual experience to judge the maple tree’s existence. But reliabilism can only work when we can specify what exactly is the adopted method and its reliability. Thus if we can’t find the single best description of a method and all

different characterizations of methods vary in their reliability, reliabilism in general is doomed. What do we do about that?

The interesting thing to notice is that the Generality problem is not a problem specifically for reliabilism, rather, it is essentially a philosophical concern about how we use the word "method". Methods allow us to achieve certain things -- securing beliefs is just one of them. Consider how we pass a ball to others. We can throw the ball or roll the ball on the ground, etc. On the one hand, we recognize "throwing the ball" as one holistic method that is different from rolling the ball. On the other hand, it is itself too general a method. There are obviously many submethods of throwing, such as "throwing with a left spin", "throwing with a right spin", "throwing with great velocity". We can see that this list can go on forever by adding extra variables such as distance, lighting conditions, or even time. If we want, we can also describe all the brain mechanisms and muscle movements as detailed as we wish. But when the description gets too specific, of course, we would be thinking that our method gets too narrow. But the question we need to ask is, do these considerations bother us in our normal measurements of the success rate of throwing a ball? When my coach says "do that again", would I be confused and have no idea what to do? Would my coach never be able to conclude that I am a good/bad ball thrower?

In contexts of ordinary language, what concerns us is only a very restricted set of variables. If my coach is testing my throwing ability in the playground, then we are not generally worried about sunlight, time, gravitational force, the wind speed, the density of air, etc. unless one or several of them is observed to have an impact on my performance. Of course, gravity contributes to my performance, but in this context, my coach doesn't really need to count it as

a variable. What is important in this case is probably my posture, the distance, and the quality of the ball. To measure my reliability in throwing the ball is to count my throwing the ball as a repeatable event. To achieve this, the coach has to hold those central variables constant, but of course, they are constant in the sense that it is appropriate for us to interpret them as such. In other words, we don't need to objectively fix these variables. It will be fine as long as these variables remain in a reasonable leeway. If, however, the wind starts blowing and intervenes my activity, then it is the responsibility for the coach to also pay attention to that variable. Our concerns for variables constitutes our interpretation of the methods.

The way we can repeat our belief-forming methods to arrive at beliefs is no different. Usually, for the visual experience, if the distance is affecting my beliefs, then obviously we need to take that into account. (This happens, for example, in an eyesight test.) But when we are just starring at stuff on the desk, we neglect the consideration of distance. Observe that how we construe meaningful variables is shaped by our previous experiences. For all we know, our visual ability is stable when it comes to close distance and declines when distance increases. If we all possess clairvoyance faculties, then distance is not a concerned variable in our method anymore. If we have extraordinary night vision, then why are we bothered by the lighting conditions so much when it comes to identifying objects? If the gravitational constant is always changing (sounds contradictory, but it's just a name), we would not have a chance to talk about weights. But when we are talking about weights on earth, we don't necessarily need to include gravitational constant. It should now be clear that when it comes to any method, there is a list of variables that we need to hold constant and a whole lot of other variables that we ignore in order for us to conceptualize a repeatable process. This list of variables is also subject to

reevaluation upon challenge. Consider the maple tree example again. When it comes to seeing, distance and lighting conditions should be our primary concerns. But our invention of the window might provide us with a third variable – transparency. When the window is transparent enough that it does not bother our visual process, we wouldn't care about this variable. On the other hand, if the window is blurry to a certain extent, then we would normally hold it constant if we want to repeat the process of seeing. This should deflect our worries about methods being too narrow. Broadness should not be a concern as well, as apparently, we know that it is the visual process that allows us to conclude colors, it is the auditory process that allows us to conclude sound, etc. Our characterization of methods should be adequate in the sense that people can act upon it. To know the area of a triangle, we can't just be told to "do the math". Unless a specific instruction is given, which should look like "measure the base; measure its height; put them into an equation and calculate', we would not have an idea of what the method is for us to evaluate. The important thing is, as long as the instructions are clear enough for us to follow, it does not really matter how we describe the method. It is not crucial at all how my coach gets me to throw the ball in the right way. He can achieve this by talking, or showing, or a combination of both. But at the end of the day, he is pointing at a single method, and he is able to tell whether I do it correctly or not. Different math teachers might explain how to use one specific technique very differently, but it does not prevent us from interpreting the technique as the very same technique. It is a difficult question as to how we take two sets of instruction to be the indicator of the same method and how we view two different sets of actions as both reflecting the same method.[28] But it shouldn't be a question that we are indeed able to

[28] We can take our impression of food preparation processes as another thought experiment. What

do so. It is undeniable that we are clustering our paths to certain kinds of beliefs into distinct categories.

Conclusion

The big question I've been asking is: Under what circumstances should we attribute knowledge to S who holds the belief P? Answering this question enables us to see how human knowledge moves around and what these activities of knowledge attribution are really about. I started with Goldman's pure externalism account of knowledge attribution together with Brandom and McDowell's critical analysis. Externalism is insufficient because committing to a belief is more than being able to express the belief itself, reason and the ability to defend cognizer's own beliefs are necessary for the game of knowledge attribution. On the other hand, externalism is right that reliability must play a central part without which the picture of knowledge cannot be fully explained. Reliability is the very concept that fundamentally explains the strength of justification and evidence. The solution is to consider knowledge as the product of the exertion of approved reliable methods. Not only does this account collects the good points from versions of externalism and internalism, but it also provides us a structured and comprehensive way to look at how human expand and edit their knowledge pool. Human communities have been constantly discovering, devising, and revising cognitive methods to secure different forms of beliefs. The methods deemed plausible are then distributed to all community members that are eligible for using those methods. This form of reliabilism isn't at odds with internalism's insights. In this picture, justification is to be understood as the indicator of the method the cognizer adopts and

---

allows us to say something like "these two chefs cook the food in the same (or different) way" when, obviously, the chefs must be performing different set of actions?

evidence is to be understood as the component of the reliable method. But one cannot use an approved method out of thin air. She must be able to legitimize and defend her confidence towards a method. Thus briefly, in a community Y, ascribing knowledge that P to a subject S is appropriate in the context C if and only if:

(1) S believes that P.

(2) S adopts a method M approved reliable by the community Y to arrive at her belief P.

(3) S is able to establish the legitimacy of using M.

To approve a method is to undertake the commitment that M is reliable. It is necessary for M to pass the provisional T/F Ratio Test, that is, it does not fail the local T/F Ratio Test using the current available data and reasons. Besides, our intuition with regards to whether a method should be approved would vary depending on the set of information available. There are no clear-cut criteria as to whether some method should be approved. We can be expect to experience this interesting indecisiveness in the waves of technology, especially in the ones associated with artificial intelligence. We are not entirely sure whether we may regard some of the algorithms as the sources for entitled beliefs. Our intuition of knowledge attribution can be shaped through constant negotiation. These novel methods for knowing encourages us to think about the process through which lines of justification become embedded in normal practices. If we then trace back to the beginning of human history, our simplest knowledge from sense perception is no different in the sense that it too had gone through this process of

social acknowledgment. Only after we figured out that our beliefs converged with that of others was observational knowledge a thing.

Knowledge isn't an infallible enterprise, nor is knowledge attribution. The approval of methods cannot escape this insecurity either.[29] Indeed, we can make moves in language games because of the reasonable assumptions we made that might be proven false. Sellars wrote, "For empirical knowledge, like its sophisticated extension, science, is rational, not because it has a foundation but because it is a self-correcting enterprise which can put any claim in jeopardy, though not all at once." (1956, 38) But even though our current beliefs are subject to future challenge, we shouldn't be worried as long as the moves we made were appropriate. To approve a method reliable is to trust the capacity of a repeatable belief-forming process based on proper evaluation, that is, based on reasons, reasonable assumptions, and the method's previous performance. Appropriate knowledge attribution then happens according to the list of approved methods. As Sellars suggested, we may find any of our beliefs along the way to be problematic and thus the associated beliefs must be reevaluated, but the whole process of assuming and revising is what allow us to move forward. This is why we should talk about appropriateness rather than truth.

[29] Here is an interesting quote from Bertrand Russel. "The demand for certainty is one which is natural to man, but is nevertheless an intellectual vice." - Bertrand Russell in "Unpopular Essays" (1950)

Works Cited

Alston, William P. "How to Think about Reliability." *Philosophical Topics* 23, no. 1, 1995, pp.1-29.

BonJour Laurence. “Externalist Theories of Knowledge”, *Midwest Studies in Philosophy,* no. 5, 1980, pp. 53–73.

Brandom, Robert. “Truth and Assertibility.” The Journal of Philosophy 73, no. 6, 1976, pp. 137-49

----- “Knowledge and the Social Articulation of the Space of Reasons.” *Philosophy and Phenomenological Research*, vol. 55, no. 4, 1995, pp. 895–908.

----- “Insights and Blindspots of Reliabilism.” *The Monist*, vol. 81, no. 3, 1998, pp. 371–392.

----- *From Empiricism to Expressivism: Brandom Reads Sellars*, Harvard University Press, 2015.

Cohen, Stewart. "Justification and Truth." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 46, no. 3, 1984, pp. 279-95.

Craig, Edward. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Clarendon, 1990.

DeRose, Keith. "Solving the Skeptical Problem." *The Philosophical Review* 104, no. 1, 1995, pp.1-52.

----- “Sosa, safety, sensitivity, and skeptical hypotheses.” In J. Greco (ed.), *Ernest Sosa: And His Critics.* Blackwell., 2004, pp. 22--41.

----- "Contextualism and Knowledge Attributions." *Philosophy and Phenomenological Research* 52, no. 4, 1992, pp. 913-29.

Descartes, René, Donald A. Cress, and René Descartes. *Discourse on Method ; and, Meditations on First Philosophy* Fourth edition. Indianapolis: Hackett Pub., 1998.

Gettier, Edmund L. "Is Justified True Belief Knowledge?" *Analysis* 23, no. 6, 1963, pp. 121-23.

Dretske, Fred I. *Seeing and Knowing* Chicago: University of Chicago Press, 1969.

----- "Epistemic Operators", *The Journal of Philosophy*, no. 67, 1970, pp. 1007–1023.

E. Conee, and R. Feldman. "The Generality Problem for Reliabilism." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, vol. 89, no. 1, 1998, pp. 1–29.

Fricker, Elizabeth. "Unreliable Testimony." In Brian McLaughlin & Hilary Kornblith (eds.), *Goldman and his Critics*. Malden, MA: Wiley-Blackwell. (2016) pp. 88-120.

Goldman, Alvin I. "What Is Justified Belief?" in G.S. Pappas (ed.), *Justification and Knowledge*, Dordrecht: Reidel, 1979, pp. 1–25.

----- "Epistemic Folkways and Scientific Epistemology." *Philosophical Issues* 3, 1993, pp.271-85.

Klagge, James & Nordmann, Alfred (eds.). *Philosophical Occasions: 1912-1951*. Hackett Publishing Company, 1993.

Lewis, David K. Elusive knowledge. *Australasian Journal of Philosophy* no.74 (4), 1996, pp.549 – 567.

McDowell, John. "Knowledge and the Internal." *Philosophy and Phenomenological Research* 55, no. 4, 1995, pp. 877-93.

----- "Knowledge and the Internal Revisited." *Philosophy and Phenomenological Research* 64,

no. 1, 2002, pp. 97-105.

Nisbett, R. and Ross, L. *Human Inference: Strategies and Shortcomings of Social Judgment*, Englewood Cliffs, N.J.: Prentice-Hall, 1980, pp. 17-42.

Nozick, Robert. *Philosophical Explanations Cambridge*, Mass: Harvard University Press, 1981.

Popper, Karl R. *The Logic of Scientific Discovery.* London: Hutchinson, 1968.

Quine, W. V. O. *Word & Object*. MIT Press, 1960.

Searle, John R. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, 1969.

Sellars, Wilfrid S. "Empiricism and the philosophy of mind. " *Minnesota Studies in the Philosophy of Science*, no. 1, 1956, pp. 253-329.

----- "Some Reflections on Language Games." Philosophy of Science 21, no. 3, 1954, pp. 204-28.

Sosa, Ernest. "How to Defeat Opposition to Moore." *Philosophical Perspectives* 13, 1999, pp. 141-53.

Williams, Michael, *Problems of Knowledge: A Critical Introduction to Epistemology.* Oxford: Oxford University Press, 2001.

Wittgenstein, Ludwig, and G. E. M. Anscombe. *Philosophical investigations* 1953.

----- *On Certainty*. Oxford: Basil Blackwell Press, 1975.