

In *Tracking Truth* I undertook a broader project than is typical today toward questions about knowledge, evidence, and scientific realism. The range of knowledge phenomena is much wider than the kind of homely examples—such as “She has a bee in her bonnet”—that are often the fare in discussions of knowledge. Scientists have knowledge gained in sophisticated and deliberate ways, and non-human animals have reflexive and rudimentary epistemic achievements that we can easily slip into calling “knowledge.” What is it about knowledge that makes it natural for us to use the same word in cases that are so vastly different? How is it possible for knowledge to have evolved? What is it about knowledge that it should enhance our power over nature, as Francis Bacon observed? What is it about evidence and knowledge that makes you more likely to have the latter when you have the former? Specialization is necessary to progress, but the division of labor it requires has allowed such questions to fall through the gaps between discussions.

These gaps are opportunities. Sometimes newly discovered problems can bring new and better answers even to old questions. The questions I have asked above are “Why?” questions expressed as (apparently) Socratic “What is?” questions, and that is the approach taken in the first five chapters of this book, to offer explanations of familiar phenomena on the basis of rigorous definitions of knowledge and evidence. One might object that this is an old, not a new, style of answer, and one that I ought to be educated enough to reject. Many have thought the project of giving necessary and sufficient conditions for knowledge was in its death rattle long ago. The most common argument for this conclusion is an empirical one, that no such attempt has ever been successful in giving the right answer for all examples. And when one asks, as one must, what the “right” answer would be anyway, the project can look even more depressing. But even if there is a clear standard we are consistently imperfect against, and always will be, I think that to view such imperfection as nothing but failure is to fall into the trap of making the perfect the enemy of the good.

It is a mindset that sadly squanders our ideas. Even if we will never succeed in the way specified, it doesn’t follow that there is nothing to be gained in the project of formulating general definitions. No scientific theory we know of has ever done a perfect, or even good, job with every case. And this is true from a theory’s inception; as Kuhn taught us, every theory is born falsified. I take it we would all hesitate, though, before concluding that this makes scientific theorizing pointless. Imperfection is our condition, but even among false theories some are better than others. We should take a big idea seriously not only, and perhaps not even primarily, in virtue of thinking it might be The One, but rather to the extent that it holds the promise of giving us more understanding than we currently have. An exclusive focus on a theory’s matching our feelings about every example we come up with also distracts us from the ultimate goal of this theorizing, which in my view is explanation. General ideas can point us to explanations that are true and illuminating in particular domains, even if as we find an idea’s boundaries it must cease to be seen as generally applicable.

I see the history of 20th century theorizing about knowledge not as converging to failure but as a sequence of ideas that were developed in enough detail to brightly illuminate the subject matter by uncovering features of knowledge we had not noticed before, some by predicting them, others by tripping over them as obstacles. Alvin Goldman’s causal theory of knowledge, for example, exposed the neglected role of the

fact that in successful perceptual knowledge our beliefs are connected to the world in a causal and historical way. (Goldman 1967) Surprisingly, causal requirements were not enough to explain why the beliefs we count as knowledge are not true accidentally, since a roll of the roulette wheel is also a causal history. However, Goldman's process reliabilism (Goldman 1979) exposed another pro-epistemic feature, reliable production of true beliefs, that both provided an explanation of why causation helps when it does, and addressed this particular accidentality problem. Similarly, Keith Lehrer and James Paxson (Lehrer and Paxson 1969) uncovered the important role of the defeasibility of evidence in our judgments of whether a belief is justified. The fact that these epistemic features had not been discussed before in traditional epistemology is sufficient to show that our understanding of knowledge has been improved, despite the fact that the causal theory, the indefeasibility theory, and even process reliabilism have important counterexamples and problems. We would understand much less than we do had it not been for the panic induced by Edmund Gettier. (Gettier 1963)

One may wonder how definitions of concepts can be explanatory when they give rise to theories that aren't true. Traditionally among epistemologists the truth of a theory of the criteria for knowledge has ultimately been measured against intuitive judgment of whether a case of belief certified as knowledge by the theory really is; we ask ourselves whether we would call that case "knowledge" if we met it on the street. One problem with this standard is that there are many cases where intuitions give no clear answer; we might, and I do, conclude that the ordinary concept of knowledge is not fully determinate. There are other cases where conflicting intuitions arise for different people, or even one and the same person; the ordinary concept of knowledge probably involves a number of overlapping but not identical heuristics that yield different answers in the non-overlapping areas. The ordinary concept of knowledge is not rigorous enough or determinate enough, and possibly not even univocal enough, to support the traditional method of evaluating theories of knowledge. And I doubt that there is a unique a priori concept hiding behind the haze, at least of the sort that has traditionally been sought. We should be less confident than people often are in the weight we ask an intuition to bear. Here, as in natural science, casual observation of cases and raw data are only part of our evidence.

This makes what it is for a theory of knowledge to be true problematic. What should it be true *to* where the data give conflicting answers, or none at all? My response is to take intuitions as defeasible constraints on theorizing, and to offer what are, quite literally, definitions, in part stipulative, for the purpose of delineating rigorous concepts that might be even more useful than the ones we encounter in ordinary language. In this way, my "What is?" questions are not Socratic, but Carnapian. (Carnap 1950, 3-8) However, though I deal in definitions, the relationships between well-defined concepts are of course factual, in whatever way logical and mathematical relations are factual, and these relationships are a key part of how definitions can be the core of explanations and keys to their discovery. If I define a simple concept like tracking, and find that its possession implies possession of another rigorously defined property that looks a lot like what Bacon described as an ability to exploit the rules that Nature herself employs (as I do – Chapters 1, 4), and if I find that what we have in what we call "knowledge" looks again and again like tracking (as I do – Chapters 1 – 4), I will have thereby offered a

plausible explanation of the truism that knowledge brings power, namely, that the thing we call “knowledge” involves tracking, and tracking brings power.

Truisms have exceptions, of course. There may be cases where something we call “knowledge” doesn’t bring power, judging intuitively. In such a case, an advocate of my theory might look for some special circumstance that explained why tracking didn’t bring power in these cases either. If this went well, then the case would count merely against Bacon’s dictum. A different kind of exception is possible, where we have a case we are happy to call “knowledge” and which seems to give power, but where tracking is absent. This would suggest the possibility of more than one kind of power that knowledge bestows and call for an attempt to identify what simple features may be at the bottom of that. In such a case, wouldn’t the tracking theory be wrong? Yes and no. The theory would have a mismatch with the data if asserted as a claim about all of it. But it would still be an explanation of knowledge’s property of power in the part of the knowledge domain where the association between tracking and power exists in the instantiations of the rigorous versions of those properties. Isn’t restricting the scope of a claim just a way of evading falsification? Isn’t theorizing in this way a friction-free and unconstrained endeavor? Definitely not; after all, the area of mismatch would have been admitted. But also, ultimately, the evaluation of a definition is pragmatic; the question about it, here, is whether it is useful in discovering explanations.

Why should a theory with mismatches be allowed to go on to live another day? Because, or to the extent that, there remains a domain in which it does correspond to and explain a correlation between properties (not only cases), and because there may yet be appropriate correlations between tracking and yet other properties. Further constraints that limit evasion come from the side of explanation. If the domain over which the relation between tracking and the intuitive notion of power shrunk to one case, for example, the concept of tracking would be useless in the explanation of power. And if the domain of real cases that instantiate the relation between tracking and power were found to overlap only slightly with cases we intuitively call knowledge, then the process of definition would have taken us to a different subject matter from the subject of knowledge that we started with. It might be a worthy subject matter, but not the one we were trying to understand. Intuitions have a role in keeping us on topic.

Since the scope of applicability of a new, rigorous definition is liable to shrink under further investigation, one might question the wisdom of even attempting to give fully general simple theories. The effort is valuable, though, because greater simplicity of a theory and greater breadth of its true implications are key features of its ability to unify our knowledge of phenomena, and unification is one key to the kind of explanations I am looking for. This focus on explanation as unification is also one key to my demand for rigor. Rigor not only makes the boundaries of the defined concept definite and more efficient to evaluate in examples, but also its often-lamented abstractness and retreat from detail bring a tendency to simplicity. The fact that the tracking conditions, to be discussed below, are profoundly simple and yet highly consequential contributes to the high potential for explanation. So too for other big ideas.

A theory of knowledge should be judged, then, not just by its tally of counterexamples, but also by its promise of explanations, and especially by the range of instances of the illuminating conceptual correlations it uncovers. It is by this standard that I found the tracking theory of knowledge to have been greatly underestimated, and saw

that further development of the view would be fruitful. I can only briefly discuss a selection of the issues here.

Nozick's original tracking theory of knowledge (Nozick 1981) focused on the following counterfactuals as requirements:

- 1) If p weren't true, then S wouldn't believe it.
- 2) If p were true, then S would believe it.

The appeal of these conditions was mainly seen at the time as immediate intuition—especially about 1)—and the nice trick that 1) played on skepticism, since it is possible to 1)-track that there is a table in front of us, without 1)-tracking that we are not brains in vats. We can thus concede something to the skeptic without any obligation to worry about our ordinary knowledge. Predictably, these conditions ran fairly quickly into a blinding snowdrift of counterexamples. These are chronicled in Chapters 2 and 3 of *TT*, as they are turned back, group by group, via two modifications of the original conditions. One is to reformulate the tracking conditions in terms of conditional probability:

- 1') $P(\neg b(p)/\neg p) > s$, where $.95 < s < 1$
- 2') $P(b(p)/p) > t$, where $.95 < t < 1$,

Where “ $b(p)$ ” says “ S believes p ,” and the thresholds s and t are determined by the disutilities of the corresponding errors in the subject's context. The other modification is to impose closure of knowledge under known implication. This means that it is possible to know p not only by tracking p , but also by knowing that p is implied by some q that you track. You have to track some statement to know p , but it need not be p . This relaxation of the tracking view is achieved by a recursion clause. It is imposed because I think, largely on the basis of arguments that have already been made, that knowledge is closed, and because though non-closure is a great move on skepticism, there's no reason to think it is more natural than closure is for a tracking theory. Objections to the brute force method of imposing the property of closure are not compelling: all the theories of knowledge we know of have to do this to get closure, as we see when formulations become rigorous. (Consider, e.g., the recursion clause needed in process reliabilism in Goldman 1979. And internalists who deny closure must think the property is independent of the concept of justification.)

Just as important as the details is the much broader appeal of the tracking idea. Condition 1, *variation*, says the knower has an ability to pick up on the matters that indicate p 's not holding, and to manage belief in p accordingly, while condition 2, *adherence*, says the knower has a sense of those matters that are relevant to p 's holding, a sense of which are not and an ability and disposition to believe in response to the former while also ignoring the latter. The idea behind both conditions is that a knower must be responsive to the world, in the midst and process of its stabilities and variations. This shows immediately why knowledge is more valuable than mere true belief: true belief is a state at a time, which tells us nothing about whether you have a chance of staying in synch with the world over time. (See Chapter 1.) The dynamical aspect of these tracking conditions also immediately explains why knowledge would give us an advantage in the attempt to exploit the rules Nature herself uses. To the extent that we are responsive to her regularities and irregularities, we must be picking up on at least some of her signals concerning p and not- p . In order to support tracking those signals must be regular

indicators of p and not- p , which means they must be kept in regular synch by nature's laws. Thus if we track our beliefs respond in accord with laws. (Chapters 1, 4.)

Responsiveness is not a vacuous abstraction, for no other theory of knowledge we have takes this direction of fit—our ability to follow the world—as essential. The world's state is the independent variable and your beliefs are the dependent variables that tracking puts conditions on. The popular property of safety—if I were to believe it, it would be true (not easily false)—involves only the opposite direction. This is inadequate, since though if I have power to make the world conform to my beliefs, that will bring safety with it, safety won't give power to me. (Multiple contrasts between safety and tracking (also known as “sensitivity”) are discussed in Chapter 4. Process reliabilism is at a less dramatic but real disadvantage here too: it says, roughly, that if I know then I used a process that will give me a true belief about p most of the time. But what if the world is such that there's a decent probability that that process I used would stop working for this purpose? We have no reassurances about that case, whereas with tracking we do, if this case is a probable enough eventuality. Internalist theories of knowledge will tend to have trouble explaining its power, since they typically require no robust relationship at all between one's belief and the world. Typically the only external requirement on knowledge for an internalist view is that the belief be true, but merely true belief may be so accidentally, and that gives us no assurances about whether believing p is the best way to manage ourselves as the world moves on and we try to work on it.

The original tracking theory had a problem with knowledge of logical and mathematical truths (as does every theory, I argue). The variation condition for these truths ends up being either undefined or trivially fulfilled, since it is not possible for these statements to be false. However, the core idea behind the tracking theory is *responsiveness*, and it makes sense that to a proposition whose truth value behaves quite differently than those of empirical propositions, we owe a different kind of responsiveness. The place where appreciation of logical truths should make us believe differently in different situations is with implication. If p implies q and someone believes p , but were he to form a belief on the matter of q , that belief would be in not- q , then we would have to say he does not know that p implies q . To know that p implies q requires having your beliefs in p and q be properly responsive *to each other*. This view also explains why it is proper to believe a logical truth itself come what may. It is not just because it will be true come what may, but because a logical truth is implied *by every proposition*. Therefore, by the account of knowledge of logical implication just described, you should be disposed to believe a logical truth r , assuming you have a belief at all, whatever your other beliefs might be. Knowledge of necessary truths is fallible—you might have been wrong—because the responsiveness you must have is not to the difference between the truth and the impossible falsity of the logically true statement, but to the relations these truths have to other propositions and to the relations they impose on the truth-values of other propositions. The required dispositions are dispositions among your beliefs, so it is both possible for you to fail, and it is possible for us to count you as knowing even if your dispositions are not perfect.

Evidence and knowledge have a relationship that, largely because of a division of labor, never gets discussed: the better our evidence for p , the more likely we are to know p . One might think this can easily be explained by someone with a probabilistic view of evidence. After all, for her e is evidence for h if e raises the probability of h . However,

that only shows that evidence makes your belief more likely to be *true*, and we've known for a long time that the truth of a belief is not enough to make it knowledge. Better evidence not only gives us a greater chance of being right, but also more of that extra robustness that knowledge has. In Chapter 5, and independently of the tracking theory of knowledge, I defend the Likelihood Ratio as the best measure of evidence, by my own and other people's arguments. This is a ratio of the probability of *e* given the truth of the hypothesis to the probability of *e* given its falsity. In other words, real evidence must discriminate between the truth and falsity of the hypothesis; the more it does this, the better it is. It will be better when the numerator is higher and the denominator is lower, that is, roughly, when *e* would allow us to better fulfill the adherence condition, and the variation condition respectively. This is the core of the argument in Chapter 5 showing a strikingly smooth connection between evidence and knowledge so understood. I go on to develop a confirmation theory that decomposes the information in evidence in a different way from standard Bayesianism, and that has advantages in allowing us to evaluate the impact of our evidence in some cases where we don't have a lot of background knowledge. It allows us to have what I call "leverage" in this and other difficult cases.

The final chapter of *TT* shows the relevance of confirmation theory to scientific realism. That is, the question what evidence is constrains our answers to the question how much knowledge of the world our evidence can and actually does get us. I argue that Bas van Fraassen's Constructive Empiricism, the best-known version of anti-realism, requires an assumption I call Equal Punishment (EP): no observational evidence can disconfirm a theory more than it disconfirms what that theory says about observables. The only plausible confirmation measure we know of that yields this result is the ratio measure (*e* confirms *h* to a higher degree the greater is the ratio $P(h/e)/P(h)$). However, the ratio measure, like all others, requires evaluation of claims about unobservables in order to ascertain whether we can legitimately make an inference from what is observed to general claims about observables. In order to avoid radical skepticism, the Constructive Empiricist claims we can do the latter. However, the former, evaluation of unobservables, is necessary for the latter, and if he says we can do the former, he has succumbed to realism. The distinction between observable/unobservable is sensible, I grant, but it's not the issue in confirmation, and can't get the anti-realist traction for a sensible middle ground between realist and skeptical positions.

The strategy anti-realists have often taken, of making claims about the limits in principle of what our evidence can confirm seem to me unwise anyway. Philosophers have regularly been embarrassed by science in our claims that one or other thing is not possible. I argue that we have actually gotten beyond the observable line, in cases like pregnancy tests, for example. However, the folly of the realist, one who thinks our best-tested theories are approximately true, has typically been to vastly overestimate how far we've gone up the ladder of confirming general theories. Using the measure most favorable to the realist point of view, the Likelihood Ratio, we see that our high-level theories have not been confirmed, due to the difficulty of evaluating the probability of the evidence given the negation of the hypothesis (the "catch-all").

Many have thought that the problem of the catch-all is one of limits on our ability to conceive alternative theories. I argue that conceivability is irrelevant to this problem, because we now are able to evaluate large classes of theories without describing each theory or even idea, or even subclass of theories, in that class. Our methods have gotten

much better over the history of science—an optimistic induction to put beside the familiar pessimistic one. So, contra the new pessimistic induction over the history of science offered by Kyle Stanford (Stanford 2006), which is based on limits of conceivability, we don't have grounds for declaring limits in principle to what new evidence may be able to confirm. The optimistic induction I proposed in the book is developed in Roush 2009.

Reply to Goldman

I would like to thank Alvin for a spirited, and gentlemanly, debate we've had on these issues, which is extended further here.

Alvin is exactly right that if we make his assumption about maximum specificity and deduceability (which I have doubts about), then on my view of knowledge Sphere Guy doesn't know there's a sphere in front of him. This may sound silly when we focus on his tactile access to the sphere in the actual world, but if we take a broader view we see that there is more at stake than this.

Contrary to Alvin's impression, methods are not at all excised from my view of knowledge. My theory of how to judge whether someone knows requires us to consider everything (probable) that is and would be responsible for the fact that the person believes or not, whether these occur in his head or in the world, which the formulation in terms of probability helps to make very clear. (See chapter 3.) Ironically, my refusal to relativize to method has us taking into consideration more facts about the method of the subject than Alvin's criteria do, for my view takes into account, as appropriate, what process the person would have used and has a tendency to use, and not just the properties of the one he happened in fact to use.

When the fact that a method was used by a subject in coming to belief in p is independent of the truth of p , which is actually most of the time in our lives, the conditions of application of the variation condition insure that we evaluate the subject by considering only what he would do and how he would fare in his beliefs were he to use that method he actually used. So, under that condition, my view agrees with Alvin, and Nozick also. But when whether a subject used that method is not independent of the truth value of p , then the variation condition in my view says we must consider in addition the subject's resulting beliefs in all scenarios where he is such that he might well have used a different method. That is, we ask about whether his tendencies in method choice are any good.¹ Relativizing a criterion of knowledge to the method the subject actually used requires us simply to ignore, in all cases, the subject's tendencies to use good methods or bad. (I will discuss this issue for the adherence condition below.)

This approach picks up other cases where intuitions are on my side. For example, do we know that we are not brains in vats when if we were brains in vats our process of belief formation would systematically hide that from us? Most people think that undermines our knowledge unless we have more to say. On Alvin's view it's not a

¹ Method is not named in the conditions of application, so there is no problem of specificity. What I say here about method is a consequence of a procedure that tells you for any q , of any level of generality, whether it gets held fixed or not.

problem: we know as long as long as we're actually using good perception. On my view we can know we're not brains in vats, but it's in virtue of tracking the table, which requires that our method of coming to a belief about that is highly unlikely to be different in case the table isn't there. If, though we are not in fact brains in vats, the scenarios in which a person like us is a b-i-v are quite likely, then this condition is not fulfilled. Though mine is like Alvin's an externalist view that does not require internal access to an argument about our situation in order for us to know, the condition I set for whether we know or not is much more in keeping with our intuitions. Do we know that there are no nuclear missiles headed for the large American cities if were there to be then it is highly likely that American intelligence would hide this from the public in order to avoid panic? I don't think we would say that we do unless we think we know, meaning track, other claims from which it follows that such a missile attack is quite unlikely.

The frequent fabricator is actually different, since which method she uses *is* independent of the truth value of *p*. Thus my variation condition relativizes to her method. Where she fails is in the adherence condition, since for that the rules of application allow anything that is independent of the truth value of *p* to vary, to see if the subject is good against their (usual) irrelevance. Here her method choice is evaluated for its general tendencies, and since she fails to have the right belief most of the time, she doesn't know.

I formulated my example in terms of frequency of fabrication because of Alvin's association of frequency with reliability. (Goldman 1979) But my point doesn't depend on this. If we don't care how it was determined that a subject would use the method she actually used, then we will allow as knowledge cases where we have unreliable usage of a reliable method or process. That makes the process of belief formation broadly unreliable. I think this issue is even more awkward than that for Alvin's view, because we can make the counterfactual question about which method to use live in the subject's head. Consider a case where the subject's brain determines, badly, whether she uses the fortune-teller or perception method. How can Alvin deny that that choice itself is a part of her process of coming to believe *p*? If it is, then Alvin must say, like me, that her belief isn't knowledge.

Alvin says I am entranced with power, and I gather I am meant to be discouraged by this. But if I'm entranced then so is the entire human species. Among his unfortunate properties Sphere Guy also has a serious survival disadvantage over his tracking counterpart: if a tiger were to come up and steal his sphere then he would likely still believe it was there because he wouldn't be using his hands anymore but his hallucination-inducing eyes. The tracker, whose method doesn't go kerflooey when the sphere goes, would pick up on the missing sphere, and if he also had reactions appropriate to the fact that tigers liked to steal spheres, he'd be more likely to take the needed measures to survive. Organisms like Sphere Guy don't survive as long as their counterparts, so if knowledge promotes survival then chalk one up for my view.

The disagreements Alvin and I have about the power-promoting property of knowledge seem to revolve mainly around how we understand the intuitive power side of the relationship. Neither of our views are developed in much detail, and I think this is a topic that deserves more attention and discussion. That said, my starting point is, of course, much better than his. Alvin says that knowing *p* at some later time does not give

the subject power at this time. I agree and haven't proposed otherwise. The question is what it is about having the property of knowledge now that gives one power now.

My view is that having power at a given time requires having a property that gives the potential for future successful action. For example, power of certain kinds possessed now pose a threat *now* (typically to others) without actually being used now. And we can say of someone who had cutting edge weapons but lost the war through incompetence that he had power but didn't use it well. That is, he had a lot of potential to win the war at the time when he started it, but as a matter of fact he didn't use it properly. If he had won the war he would have exhibited the power that he had, that is, the potential that existed from the beginning. Having power doesn't determine what you will do in the future, but power at a given time enhances your potential for future action.

The having of a functioning memory now will of course also be part of what supports your potential now for acting in the future, but this is all included in tracking the world now. If you don't have a functioning memory now you are not now prepared in your belief dispositions for all those possible scenarios in which you would have to decide whether to change or stay with your beliefs. By contrast, having as a matter of fact formed the belief that you have now through a reliable process doesn't tell us what you now have a tendency to do with your beliefs in scenarios that you are not now actually in. One reason is that for knowledge now Alvin doesn't require that the subject have a tendency to use a good method of belief formation, but only that he actually formed the current belief using a good method. So, I think that Alvin's kind of knowledge does not pose a threat to mine.

Concerning probability, my claim, in context, was that the frequency interpretation is no more controversial than the propensity or other objective interpretations. All interpretations of the axioms have their advantages and drawbacks. I tend to write using objective language for ease of understanding. As I said in the book, the definition of knowledge I am pushing is independent of the interpretation of probability; that is, it can be of service using any interpretation you like. You do get a different tracking theory depending on which interpretation you use. On an objective interpretation, of course, you get a theory of the conditions under which a subject knows. If you apply the theory using a subjective interpretation then yours will be a theory of which propositions the evaluator's belief in which should determine his belief about whether the subject in question knows. My view is that either approach should, or at least is able, to incorporate tracking. Incidentally, the interpretation of probability has nothing to do with externalism or internalism, which is about whether those propositions in the language that concern the subject's relation to the external world are relevant to whether he knows. The question of which interpretation of probability to use is instead about whether those claims are claims of fact or beliefs of the evaluator. An internalist is welcome to say that facts, rather than our beliefs, about a subject's thoughts are the only things relevant to whether he in fact knows. On the other side, even using a subjective interpretation of probability the tracking view is externalist.

The rules of application for the tracking conditions make them more determinate, yet still independent of the interpretation of probability. Though their formulation must be complicated, they express two simple ideas. 1) We determine which propositions at which level of generality are to be used in the evaluation of the tracking conditions by starting with the *p* of "S knows *p*," which itself has a particular level of generality. For

any q of the language, of any level of generality, we plug q into a rule, and out pops an answer about the role it will play – fixed or varying – in each of the tracking conditions. This is a solution to the generality problem for tracking, but it is a strategy that is not available to the process reliabilist since his criterion begins with the process rather than the belief. 2) The second simple idea is that those propositions whose probabilities should be allowed to vary when we ask what the subject's beliefs do over a variety of situations, as we do in both tracking conditions, are the ones that are more "influenced" by the truth-value of p than the truth value of p is "influenced" by their truth values. The subject should be able to "catch" the indications that p is true and ignore others that are irrelevant, and should be able to catch the indicators that p is false and ignore the others. The indicators are simply the matters "influenced" by the truth value of p .

Finally, all of these rules are relative to not only an interpretation of probability in general—choose the one you like—but also to a particular probability model, that is, a language, and a function from every proposition in that language to a value that is its probability. I call this function P_u , the Ur-probability function, and it is given by one who applies this theory, not by me. This is innocuous; it corresponds to the common practice and assumption that we can expect no answer to the question whether S knows p until someone supplies a description of S , the world, and what S 's actions, tendencies, history, and circumstances are in the world, and it is not the theorist's job to do that but the job of someone who applies or tests the theory. The function P_u simply incorporates the evaluator's complete description of the world and the subject, so that these assumptions are available as needed.

Response to Godfrey-Smith

Thanks to Peter for his remarks. It is always a pleasure.

Peter helpfully identifies three kinds of discussion of knowledge, all of which I approve of, 1) investigation of ordinary peoples' use of ideas and words like "knowledge," 2) how humans and other animals are actually connected to the world, and 3) how those concepts and our dealings with the world do or do not relate. And I agree that finding harmonies tends to show us something useful, and that this is one of the things I've been up to. But I also look for stipulative definitions that carve into the blurriness we get at the level of ordinary talk, which means I'm willing to be more prescriptive if I find that a crisp concept identifies a phenomenon that we are interested in. In other words, I'm willing to be more dictatorial than Peter may be comfortable with.

Peter is skeptical that the adherence condition is reflected in the ordinary concept of knowledge. I would still defend my account of one route to knowledge that Betsy has in the face of Jesse James, since my view of knowledge of implication is different from most in that I do *not* require the subject to even be in any position to make an argument. Although there are plenty of cases where one needs to have an argument to get to knowledge because the implication is complex, that is a contingent matter concerning the means to knowledge, not a matter of the definition of the goal; in simple cases one who knows has the appropriate responsiveness (described above) automatically. If Betsy doesn't have these dispositions between all she knows about the legendary Jess James

and the identity of the face that has been presented to her, then she is more impaired than the example assumes.

However, after the book was published I discovered that on my view those subjects who use a method accidentally also have a more direct way of knowing. For in cases of accidental use of a method, the scenarios where the accident didn't happen are ones where the subject doesn't form a belief at all. Such scenarios are knocked out of consideration in the first domain-limitation in the rule of application for the adherence condition. The adherence conditions thus ask: If she forms a belief at all, does she form the right one? Accidentality of method choice is not the problem for me that it was for Nozick.

Peter is right to point out that the president I described is faulty in too many ways to be probative here. Consider instead the Underconfident Student. In a particular subject matter she's well aware, perhaps too aware, of everything she's ever been taught about the ways that a claim in that area can be wrong, so, let's say, she would never believe or assert a p in that subject matter if it were false. However, partly because of this awareness, she often has trouble bringing herself to confident beliefs. But what if we find her on a day of high spirits; she's confident in a seminar, where she makes a presentation arguing very coherently for p , where p is true. But suppose that an aggressive fellow student were to hold forth insistently against p or her arguments for it, by talking about a matter that used similar words but was only tangentially related to p or the argument. In the case of this bluster, which is all too realistic, the person we are imagining could easily lose faith. Maybe we should count someone with this pattern of behavior as knowing in those rare cases where she in fact confidently and legitimately asserts p , but statistics and experience tell us that we actually usually don't. We expect a little rigidity—even in the face of bluster—that is based on a little confidence that such remarks may not be relevant. We expect some degree of adherence.

The Jamesian connection, the Classical Statistics connection, and the signal detection connection to the tracking conditions are all highly appropriate and welcome. The trade-offs between the two conditions are especially worthy of further discussion. Though the contingent conditions organisms find themselves in clearly have a trade-off structure due to finite resources, I have wondered whether the trade-off is due only to empirical facts or also to something about the abstract relation between the tracking conditions. And it turns out that the tracking conditions do impose a trade-off in so far as they place requirements on what you do when p is true and when p is false, and those happen to be mutually exclusive and exhaustive of the ways the world might be with respect to p . So, in principle there's no avoiding a trade-off between them. However, in practice the more resources we have, the more separating indicators we might find, which would increase the distance between the distributions, and which perfectly matches my view of what evidence is. (Chapter 5)

As Peter points out, our choices when confronted with the curves are to use the indicators we have or expend resources to find more discriminating ones. One instance of this is the choices governments make as to whether to fund more science or to apply the markers we have already found. Scientists are making the effort—and a resource-intensive effort it is—to go off the charts to find more indicators of how the world is. An attempt to balance the trade-off between the two kinds of error can also be seen in a division of labor within intellectual communities along a spectrum from stubborn

dreamers who do not want to miss any truth to those who are at least as concerned that we avoid wasting our time on nonsense.

Carnap, Rudolf (1950), *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Gettier, Edmund (1963), "Is Justified True Belief Knowledge?" *Analysis* 23:121-3.

Goldman, Alvin (1967), "A Causal Theory of Knowing," *Journal of Philosophy* 64: 357-72.

----- (1979), "What is Justified Belief?" in G.S. Pappas (ed.), *Justification and Knowledge*. Dordrecht: D. Reidel, 1-23.

Lehrer, Keith and Thomas Paxson, Jr. (1969), "*Journal of Philosophy* 66: 225-37.

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

Stanford, P. Kyle (2006), *Exceeding our Grasp: Science, History, and the Problem of Unconceived Alternatives*. New York: Oxford University Press.

Roush, Sherrilyn (2005), *Tracking Truth: Knowledge, Evidence, and Science*. Oxford: Oxford University Press.

----- (2009), "Optimism about the Pessimistic Induction," *New Waves in Philosophy of Science*. Palgrave MacMillan.