THE EMPIRICAL CONTENT OF THE EPISTEMIC ASYMMETRY

DOUGLAS KUTACH

Abstract. I conduct an empirical analysis of the temporally asymmetric character of our epistemic access to the world by providing an experimental schema whose results represent the core empirical content of the epistemic asymmetry. I augment this empirical content by formulating a gedanken experiment. This second experiment cannot be conducted using any technology that is likely to be developed in the foreseeable future, but the expected results help us to state an important constraint on our ability to know what will happen in the future. Finally, I show that a third experiment concerning precognition, described by Michael Scriven and John Mackie, does not characterize any additional empirical content but does help to illustrate why it is unlikely that any precognition exists.

The epistemic asymmetry consists in the apparent fact that our knowledge of past events is typically more secure, more detailed, and more abundant than our knowledge of future events. One efficient way to explore the underpinnings of the epistemic asymmetry is to apply the method of empirical analysis (Kutach 2010, 2013). Empirical analysis is the activity of formulating and systematizing concepts in order to optimize explanations of empirical phenomena, especially insofar as they are characterized in terms of experimental schemas. In general, the goal of an empirical analysis of some concept X is to engineer a conceptual framework that facilitates an explanation of the empirical phenomena that make the concept X worth having. The goal of an empirical analysis of knowledge, for example, is to develop concepts honed for explaining experiments that reveal our epistemic abilities.

Empirical analysis differs from orthodox conceptual analysis in two ways. First, it rejects the dogma that a conceptual analysis of X has a mark against its adequacy when there exists a common sense opinion that conflicts with the completed analysis. For example, the legendary analysis that says a belief counts as knowledge if and only if it is justified and true, is commonly believed to have been refuted by the Gettier counterexamples, scenarios many people instinctively

Date: July 18, 2014.

classify as exhibiting a lack of knowledge. In an empirical analysis of knowledge, such counterexamples are irrelevant because the goal is to construct a scientific theory of knowledge, not a theory of our folk psychology of knowledge. If it proves convenient to count Gettier cases as instances of knowledge, an empirical analysis should do so. The intuition that they fail to exhibit knowledge can be cashed out in terms of a theory of why it is handy for people to have an instinctive knowledge-attribution capacity that respects the Gettier intuitions.

Second, an empirical analysis of X *guides* us away from the platitudes that we think characterize X by pressing us to formulate experiments whose results capture why it is reasonable for people to have a concept of X.

What follows is an attempt to identify the empirical content of the epistemic asymmetry by formulating an experimental schema—the *basic epistemic asymmetry experiment*—that helps to organize the empirical phenomena that need to be addressed by any adequate empirical analysis of the epistemic asymmetry. I will explore two additional proposals in order to ascertain whether they can be formulated adequately as experiments and, if so, whether the expected experimental results constitute a dataset that goes beyond the results of the basic epistemic asymmetry experiment. The first of these two proposals does succeed in going beyond the basic epistemic asymmetry experiment and provides a useful clarification of the limits on our ability to predict the future. The second, when suitably clarified, merely recapitulates the empirical content identified previously. I cannot prove that the two experimental schemas I identify encompass all the empirical phenomena bearing on the epistemic asymmetry, but I suspect this conjecture is not far from the truth.

The philosophical upshot is that after we are in a position to provide an adequate explanation of the experiments that capture the empirical content of the epistemic asymmetry, the empirical analysis is complete. Although the ultimate purpose of any empirical analysis is to facilitate scientific explanations, I do not have the space here to provide a comprehensive explanation of the epistemic asymmetry. The goal of this paper is to identify what an adequate account of the epistemic asymmetry needs to explain.

1. The Basic Epistemic Asymmetry Experiment

When conducting an empirical analysis of knowledge, one needs to formulate an operational procedure for assessing knowledge. The most straightforward way to test an agent's propositional knowledge is something like the following:

(1) Ask the agent whether P.
(2) Receive a response.
(3) Independently assess whether P.
(4) Conclude that the agent knows P iff the response matches the independent assessment.

Students are routinely tested this way, and we generally view such methods as providing a reasonable (though certainly imperfect and perhaps very misleading) measure of how much each student knows about the tested subject matter. The crudity of such a test can be mitigated by addressing its many limitations.

First, such a test is effective only if P's truth value can be reliably determined without too much controversy. It is useless to test agents on whether their favorite color is undervalued by society or whether Gauss was smarter than Einstein. Fortunately, this limitation does not hinder an investigation of the epistemic asymmetry because the epistemic asymmetry can be stated solely by reference to public events that are uncontroversially observable. Among all publicly observable events, people know much more about the past ones than the future ones.

Second, such a test is incapable in any single case of distinguishing between an agent who has a firm grasp of the correct answer from an agent who just guesses correctly or has acquired the correct belief by luck. These limitations might be viewed as inadequacies, but they are actually a blessing in the sense that once we have adopted an externalist approach towards knowledge, which is appropriate for empirical analysis, several controversial epistemological controversies that are not addressable by empirical analysis no longer need to be settled. In an empirical analysis, we do not need to adjudicate whether an agent had that extra something that made her true belief bona fide knowledge. Furthermore, the role of luck can be mitigated simply by testing large numbers of agents to find patterns that reliably appear in large samples.

Third, such a test by itself is unable to distinguish agents who do not know the answer from agents who misunderstand the question or deliberately attempt to give the wrong answer or make mistakes in converting their knowledge into a response. These limitations can be addressed by conducting further tests that diagnose miscommunication and by providing each subject with a cash reward for getting answers right.

A significant benefit of this operational procedure for assessing knowledge is that it can be extended to check a variety of epistemic claims. We can test how a person's knowledge correlates with various factors such as household income and formal education. It extends easily to become a test of know-how rather than know-that by asking the agent to perform a target task. It also encompasses the quasi-epistemic abilities of entities we do not normally categorize as intentional agents. For example, we can test a computer to see whether it "knows" the local price of bananas.

Building on this basic idea, one can arrive at a more precisely specified experimental schema that I think fairly reasonably characterizes the core empirical content of the epistemic asymmetry. This *basic epistemic asymmetry experiment* tests how reliably an agent can judge whether an event of some chosen event-type E occurs, where the crucial independent variable is the time interval between (1) a target space-time location where E will either occur or not and (2) a deadline before which the agent must decide whether E occurs.

- For the experiment as a whole, there exists a single description of a coarse-grained event, E, called the 'target event', which is intended to represent a type of event whose occurrence is in principle empirically accessible. An arbitrary reference point t is chosen inside the block of space-time where the target E might occur.
- For each experimental run, there must be at least one agent that exists within the fragment of history constituting the experimental run. Exactly one agent counts as *the* agent of the experimental run.
- At the beginning of each experimental run, each agent is made aware of the description of the target event, E. In effect, each agent is being asked the question, "Does E occur (in your fragment of space-time)?"
- At the beginning of each experimental run, each agent is randomly assigned a deadline corresponding to a random temporal duration Δt, which can be positive, negative, or zero. The deadline is the moment that occurs Δt units of time after t; more technically, it is a space-like hypersurface. For example, E could be "There are more than five thousand dolphins alive on Earth during the first month of the year 2040," in which case t could be set as the space-time point at the middle of the Earth at the beginning of the year 2040. If Δt happens to be -10 years, then the deadline is exactly 10 years before t. If

Δt happens to be +5 months, then the deadline is exactly 5 months after t.
- Each agent is free to declare a single "yes" or "no" answer at any time before the deadline.
- The agent's answer is correct iff (1) the answer is "Yes" and E occurs or (2) the answer is "No" and E does not occur. Correctness is assessed by the experimenter using reliable independent means to establish whether E occurred.
- Each agent is given a sack of gold iff his answer is correct.
- The agents are aware of all these rules to the extent they can be aware of them.

To run some version of the basic epistemic asymmetry experiment, an experimenter would recruit a zillion agents, select a target event-kind E, inform the agents of E, and then randomly select a zillion durations, Δt_a, ranging from negative to positive, and communicate each value to its assigned agent, a. Then, the experimenter would check for any response the agent gives before the deadline, and check to see whether E occurred. This provides extensive data concerning the epistemic accuracy of agents as a function of the agent's deadline relative to E's occurrence or non-occurrence.

The results I would predict for various implementations of the epistemic asymmetry test are difficult to state in precise terms, but one can summarize the empirical content of the epistemic asymmetry as follows:

> There are *many* target event-kinds, E_i, such that agents succeed much more often when the deadline occurs significantly after E_i than when the deadline occurs significantly before E_i, but there are *few* target event-kinds, E_j, such that agents succeed much more often when the deadline occurs significantly before E_j than when it occurs significantly after E_j.

Although imprecise, I think the empirical phenomena summarized by this statement constitute the primary reason we are justified in thinking of the past as being epistemically more accessible than the future.

This crude characterization can be improved along several dimensions by taking into account various limitations of the experimental design. For instance, nothing in the experiment distinguishes between events an agent could predict correctly if there were enough time to perform calculations and events an agent is altogether not in a position to predict reliably. An agent might perform well at predicting tomorrow's weather if given two weeks for computation, and this could be accommodated by adjusting the experiment so that at the deadline the agent is not allowed to gather any more empirical data but is given extra time for calculations.

An important distinction was drawn by Dummett (1964) between two kinds of knowledge about the future: prediction based on causal laws and knowledge in intention. His idea is that if an agent is assigned to predict some future target event he has significant control over, such as whether someone nearby will clap within the next five seconds, the agent could simply predict that someone will clap and then himself clap to make the prediction successful. Fortunately, in most cases, the above experiment can be modified to distinguish between these two kinds of knowledge by running further experiments where the agent is isolated or restrained after making the prediction. Or one could run separate experiments to assess how effective agents can be at promoting the occurrence of the target event within the allotted time, by randomly assigning them the task of bringing about the target event E or preventing E, and then using those results to adjust the conclusions drawn from the basic epistemic asymmetry experiment.

One irritating aspect of the results of the basic epistemic asymmetry experiment is that they are not amenable to a convenient summary that universally quantifies over a broad class of clearly definable event-kinds. Some kinds of events are hard to predict but easy to observe retrospectively, like the outcomes of lotteries. Other kinds of events are just about as easy to predict beforehand as they are to observe afterwards, for example that the moon will exist tomorrow. The irritation is that event-types exist largely on a continuum without any sharp distinction between the kinds of events that exhibit the epistemic asymmetry and those that do not. What's more, there is no sensible way to quantify event-kinds in order to clarify what it means for "many" target event-kinds to be more knowable afterward than beforehand. Nevertheless, the core empirical content of the epistemic asymmetry can be adequately captured in the form of an infinite set of target event-kinds E_i, each of which is mapped to the results one gets (or would get) by running zillions of experiments using E_i.

2. The Prediction Experiment

One proposal to circumvent the frustratingly vague reference to "many" event-kinds was provided by David Albert (2000) when he claimed (p. 122) that what "distinguishes the sort of epistemic access we have to the past from the sort of epistemic access we have to the future" is the following pair of principles, which I will refer to collectively as (A):

> Start with a probability distribution which is uniform—on the standard measure—over the world's present macrocondition. Conditionalize that distribution on all we take ourselves to know of the world's entire macroscopic past history. . . . Then evolve this conditionalized present-distribution, by means of the equations of motion, into the future.
>
> This will yield (among other information) everything we take ourselves to know of the future.
>
> Conversely:
>
> Start with the same uniform probability-distribution over the present macrocondition. Conditionalize this distribution on everything we take ourselves to know of the world's entire macroscopic *future* history. . . . Then evolve *this* conditionalized present-distribution, by means of the equations of motion, into the *past*.
>
> This will yield immensely *less* than we take ourselves to know of the past.

It is difficult to provide a suitable evaluation of (A) because it is unclear what framework is appropriate for judging its adequacy. Albert claims that (A) is what "distinguishes" the asymmetry in our epistemic access, but in what sense? Is (A) intended merely to characterize one prominent asymmetrical aspect of our epistemic access (and so provide one way to distinguish between knowledge of the future and knowledge of the past)? Or is (A) intended as a comprehensive characterization of the asymmetry of our epistemic access, so that all existing differences between our knowledge of the past and our knowledge of the future are entailed by (A) in conjunction with innocuous auxiliary premises? It is difficult to believe it could be a comprehensive principle because all it says about our knowledge of the past is that it exceeds our predictive knowledge of the future, and that claim is compatible with an extremely wide range of possibilities, many of which we are in an excellent position to rule out. For example, we can rule out the possibility that we have only one or two items of knowledge about the past that exceed what we can infer using Albert's inference procedure, and we can rule out that we are omniscient about the past.

Furthermore, the content of (A) itself is underspecified from the perspective of empirical analysis. Is (A) intended as an empirical claim? If so, what kind of data would falsify it? Is (A) intended to *explain* some empirical data, or does it merely summarize (or constitute) the data to be explained? (A) purports to connect what "we take ourselves to know" about the future with what a certain idealized inference procedure generates from the present macrocondition and what "we take ourselves to know" about the past, but what does "take ourselves to know" mean operationally? I am afraid to hazard a guess.

In order to further the discussion, I will attempt in the rest of this section to establish that a proposition somewhere in the neighborhood of (A) constitutes an important supplement to the core empirical content of the epistemic asymmetry, and I will attempt to specify a thought experiment whose expected results would (if the experiment could be conducted) constitute the supplementary empirical content. So, using (A) as inspiration, I will now identify some empirical phenomena not captured by the basic epistemic asymmetry experiment.

The *prediction experiment* involves agents gambling on whether a chosen target event E will occur in the future. All of the rules defining the basic epistemic asymmetry experiment hold except for the following changes.

- Instead of having the randomly assigned deadline vary among agents, the same deadline is assigned to all agents, and it always occurs before E.
- There is a designated predictor in each experimental run. The designated predictor is an agent (very liberally construed) who gathers as much information as possible about the laws of nature and about the material layout of the world at the deadline. The designated predictor is prevented from receiving any information after the deadline, and it is required to specify its information in terms of a state description that solely concerns what is happening at the deadline. The designated predictor then uses the information about the deadline together with any information it has about any laws or causal principles to calculate a probability p for E. It submits p as its answer.
- Each agent will be rewarded or punished according to the following rule:
  - If E occurs and the agent correctly answered "Yes," then the agent is paid $\$999 \times (1-p)$.
  - If E does not occur and the agent correctly answered "No," then the agent is paid $\$999 \times p$.
  - If E occurs and the agent wrongly answered "No" or did not answer, then the agent must pay $\$1001 \times (1-p)$.
  - If E does not occur and the agent wrongly answered "Yes" or did not answer, then the agent must pay $\$1001 \times p$.

The payment scheme is analogous to those used at any casino. Players are rewarded/punished in the long term for how well they outperform/underperform what would be expected on the basis of chance, taking into account that the payouts include money for the casino. The prediction experiment rewards/punishes agents in the long term for how well they outperform/underperform the designated predictor's assessment of E's probability, including the modest fee.

When conducting a prediction experiment for a target event-kind E, one recruits a gazillion different agents with each one enduring a bajillion experimental runs.

The prediction experiment can be thought of in its ideal form as a mere gedanken experiment, or it can be thought of as a real experiment using the best technology currently available to create as good a designated predictor as possible. Let us consider these in turn.

First, in its ideal form, the designated predictor knows S, the complete microstate at the deadline, and L, the fundamental laws. The designated predictor is also able to ascertain anything implied by $S\&L$. Some paradigm fundamental theories postulate abundant relations of determination between fully specified time slices at different times. Other paradigm fundamental theories only use relations of determination as a default rule for temporal evolution that is occasionally overridden by injections of fundamental chanciness. If we assume there are fundamental laws that resemble these paradigm laws of fundamental physics, then $S\&L$ will imply a probability distribution over all possible ways that S could evolve into the future. Hence, S will fix determinate probabilities for all events in its future, including the target E. If the laws permit sources of indeterminism other than fundamental chances, such as naked singularities or Cartesian souls, then the ideal predictor might not be able to ascertain a probability for E. For the sake of discussion, however, I will set aside such possibilities and assume that any actual complete microstate S fixes a determinate probability for everything that could happen in the future.

My prediction for the idealized prediction experiment is that regardless of the target event-kind and regardless of what kind of agents are used as subjects, after a bajillion experimental runs for a gazillion different agents, virtually no agents will have won any money, and those that have won have done so merely by being fantastically lucky. Assuming my prediction for this gedanken experiment is correct, there is a simple (and arguably simplistic) explanation available. If the fundamental laws are not chancy, then it is mathematically impossible for the agent to win any money. If the fundamental laws include some chanciness, then an agent can win systematically only by exploiting some reliable pattern in the future layout of matter concerning events of type E that goes beyond the patterns that ought to be expected from the fundamental chances. A plausible physical principle (if the fundamental laws of the actual world are chancy) is that chance outcomes do not exhibit any reliably predictable further structure. Stating this principle in more precise terms is an important and subtle philosophical exercise. On the one hand, we need to allow that the actual microscopic history (under the assumption of extensive chanciness) is itself extremely improbable and to allow that a retrospective examination of its contents will reveal many surprising and improbable coincidences. On the other hand, the total collection of surprising patterns found in retrospect presumably does not greatly exceed the number of surprising patterns one would expect given that the world evolved by chance. We can attach a label to this supposed fact about the world by saying that the world's evolution is *future-typical*. I suspect that this future-typicality plays a crucial role in explaining the epistemic asymmetry, and to the extent that deeper explanations are available, they work in part by explaining why future-typicality holds.

Second, in a more realistic form of the prediction experiment, the designated predictor can be thought of as a machine employing a vast array of detectors spread throughout the environment to get as accurate a reading as possible on the state of the world at the deadline. This data is assembled into a representation of the deadline state, $\overline{C}$, defined as a probability distribution over a set of possible microstates occupying the space-like hypersurface that constitutes the deadline. Then, a supercomputer computes the lawful evolution of a sufficiently representative subset of the microstates in $\overline{C}$ towards the future until each one reaches the location of the target event E. The calculated probability for E is just the proportion of such propagated microstates that instantiates E in the target location. This value is submitted as the designated predictor's final answer, p.

The reason the experimental design requires the designated predictor to assemble its empirical data first into a claim that only concerns what is happening at the deadline is to reduce the possibility of its smuggling information previously obtained by precognizers. Because there is no restriction on the kind of system that can serve as a designated predictor, if there were crystal balls that could accurately reveal future events, then a designated predictor could just include a crystal ball among its detectors and thereby predict whatever is showing up in the crystal ball and thereby undermine our ability to establish a good case for the existence of precognition by consistently winning money in the prediction experiment. By requiring designated predictors to use only their estimate of what is happening on the deadline hyper-surface to propagate towards the future using their best guess at the dynamical laws and other causal principles, the incorporation of precognitive resources in the designated predictor is mostly eliminated. The loophole can be completely closed, I suspect, by also forbidding the designated predictor (when it is propagating $\overline{C}$ into the local future) from propagating $\overline{C}$ so far as to create a temporal loop that includes $\overline{C}$. That restriction prevents the designated predictor from picking up any extra information about the future by exploiting the severe consistency constraints that appear when a fundamental nomic connection hops back in time though a wormhole.

For realistic prediction experiments (conducted within the next few millennia, at least) using designated predictors that attempt to approximate the fundamental laws, I predict that for some events, there will be plenty of ordinary human agents who will win a fortune. After all, current technology is such that it is difficult for a designated predictor to acquire sufficient detailed data about what is happening at the deadline, and it is difficult to propagate microstates accurately into the future using current algorithms and computer hardware. For example, humans can accurately predict that the Onandaga Brood of cicadas will reappear every 17 years, but it would be fantastically difficult to make the same prediction by simulating the motions of all particles in the solar system for decades.

There is some conceivable level of technology such that once the designated predictor has that technology, it can approximate the predictions of the ideal predictor. In that case, it will become more difficult for agents to win any money in the prediction experiment. A residual limitation of these nearly ideal prediction experiments, however, is that if there are any agents who make money in the long term by what appears to be a precognitive ability, there will be no sure way (that I can see) to verify that a noteworthy precognitive skill is operative rather than merely a deficiency in the designated predictor's model of $\overline{C}$ or the laws.

Setting aside these practical difficulties, I think there is something to be learned from a conceivable designated predictor that knows all the fundamental laws and has an imperfect but extremely accurate representation, $\overline{C}$, of the microstate at the deadline. Previously, with the ideal predictor, there was a divergence in the explanation of the experimental result that all agents lose in the long run. If the deadline microstate *determines* E, the agents are guaranteed never to win, but if the laws include fundamental chanciness, it is the future-typicality of the actual world that explains their ever mounting losses. Now, a nearly ideal designated predictor will need to have probabilistic predictions owing to uncertainty in its measurement of the deadline state. A claim I would like to defend—though it would require far more attention than I can give it here—is that once we incorporate a very slight degree of fuzziness in the deadline state via the probability distribution built into $\overline{C}$, we will get non-trivial probabilities for future events that can mimic what we get from fundamental laws that generate pervasive chanciness. Furthermore, I believe it is reasonable to expect that once there is a little bit of fuzziness in $\overline{C}$, there will be quite a bit of insensitivity to the precise probability distribution used in $\overline{C}$. For ordinary target event-types, it would not make too much difference if the probability of a quark here or an electron there were adjusted. There would certainly need to be some constraints on the probability distribution—for example, to represent the strikingly different statistical behavior of fermions and bosons—but we would not need anything as strong as a unique objective probability distribution to serve as the one true way to represent the designated predictor's uncertainty about the deadline state.

Assuming the existence of this sort of non-fundamental chanciness, we can provide an explanation of the experimental results that is insulated from the difficult-to-answer question of whether there are fundamentally chancy laws. Whether chanciness is fundamental or merely derivative (in virtue of a fairly chaotic deterministic evolution of the physical state), it is the future-typicality of the universe that is responsible for agents being unable to predict the future better than the nearly ideal predictor. Again, future-typicality needs to be spelled out in greater detail, but I hope this brief account begins to focus the structure of the explanation.

Even without knowing precisely how to formalize future-typicality, we can recognize that the evolution of the actual microstate towards the past is highly atypical because it involves all sorts of seeming coincidences in the motion of particles. This past-atypicality, we may safely surmise, plays some role in an explanation of why we have greater epistemic access to the past. If we were to run the prediction experiment as a retrodiction experiment, asking agents to answer at the deadline whether a previous E occurred, they would easily win money against a designated predictor that was inferring the past entirely from its representation of the state at the deadline (without assuming the low entropy of the early universe). Because of the fantastic sensitivity of the actual microstate under its development towards the past, even a very nearly ideal predictor of the future would almost certainly fail to be even a moderately good retrodictor of the past. These expected results vindicate the second part of Albert's proposal.

To relate this discussion back to the first part of Albert's proposal, the conclusion I draw is that (A) turns out to be on target as a characterization of one remarkable aspect of the epistemic asymmetry, but it is stronger than necessary. Albert's proposal that the limit on our ability to predict the future is set by the "probability distribution which is uniform—on the standard measure—over the world's present macrocondition" can be improved, I think, by replacing it with "what is implied by the world's laws and present microcondition." This has the effect of elevating the standard for an ideal prediction, and that standard can be stated without saying anything whatsoever about statistical mechanics. Because any actual agent is far more ignorant about the present macrocondition than the nearly ideal designated predictors and is far more ignorant about the present microcondition than the ideal predictor, it does no harm—and greatly simplifies the overall account—just to drop the privileged status of the statistical mechanical probability distribution and instead incorporate any relevant statistical mechanical principles governing probability distributions into an account of the discrepancies that exist between ideal predictors and nearly ideal predictors. This approach allows us to avoid adopting Albert's contentious claim that there is only one correct probability distribution for predictive inferences and also allows us to avoid drawing any principled distinction between macroscopic and microscopic. In this way, we can avoid having to grant Albert's model for statistical inference a special status in our account of epistemic access and instead think of it (or something like it) as just one of many ways non-ideal inference can be modeled. One can perhaps order various models of inference in terms of their degree of idealization, starting with a model of ideal epistemic access like the one I described, and proceeding through other models like Albert's that demand less of the designated predictor but are still predicated on using fundamental laws for the inference, eventually reaching models of inference that are psychologically realistic but make no use of fundamental laws.

In summary, two data sets relevant to the epistemic asymmetry have now been identified: the actual results of various implementations of the basic epistemic asymmetry experiment and the expected results of various idealized versions of the prediction experiment. I suspect we believe in the epistemic asymmetry almost entirely because we are aware of regularities in the first data set, and that is why I dubbed it 'the core empirical content'. However, even though nothing in our experience comes close to an implementation of the prediction experiment, I suspect that the wise are aware of the most salient pattern in the second data set, which is that many identifiable kinds of observable processes are reliably and stubbornly chancy (even if fundamentally determined) in a way that apparently cannot be defeated by any training in the psychic arts or by technological advances.

By running the basic epistemic asymmetry experiment, we would not only uncover specific patterns of epistemic asymmetry associated with particular event-kinds, E_i. We would also uncover a remarkable general pattern that can be *roughly* characterized as follows: Among the various E_i that exhibit a strong temporal asymmetry—meaning a high probability of being known afterward but a low probability of being known beforehand—there are no cases where a special subclass of agents (or agents aided with special technology) are able to predict reliably beforehand whether E_i will occur. That is, there are no oracles or crystal balls or anything else that would count as an exception to the general trend that E_i is known much better in retrospect. It is the purpose of the prediction experiment to play a supplementary role by refining this pattern into a more precise characterization of the boundary between prediction and precognition.

For the purpose of an empirical analysis of the epistemic asymmetry, the upshot is that if these two experimental schemas capture the empirical phenomena that motivate our belief in the epistemic asymmetry, then one should seek a scientific explanation of their results in a way that fits comfortably with the rest of science. According to the method of empirical analysis, no further explanation is needed to account for the epistemic asymmetry. In particular, there is no need to wait for epistemologists to figure out what counts as knowledge in order to understand why we know more about the past.

3. The Precognition Experiment

In order to clarify the character of precognition, I will now explore a third kind of experiment called the *precognition experiment*. I believe that the precognition experiment turns out to be merely a variation on the basic epistemic asymmetry experiment, but it draws our attention to two distinct ways the world could instantiate precognition—causal and non-causal—and that in turn helps to illuminate several components of our naive grasp of precognition.

Michael Scriven (1956) and John Mackie (1973) offer versions of an experiment attributed to Whately Carington, which would provide evidence for precognition if certain positive results obtained. Human agents are asked to complete a drawing by Monday of a picture that will be created Tuesday using a random picture-generating device. The agents are prevented from manipulating the device (in the rough-and-ready pre-theoretical sense of 'manipulation'). The original picture hangs in a designated room throughout Tuesday evening and is destroyed at midnight. In Scriven's version, the picture is selected through a roundabout random selection of an appropriate noun from a dictionary. If a remarkable statistical correlation obtains between the drawings and the pictures after numerous repeated runs, that is prima facie evidence for precognition.

The precognition experiment appears to be a modified version of the basic epistemic asymmetry experiment in the following sense. The target event E is the full history of the randomly generated picture, and the agents are asked to respond to "What pattern does E instantiate?" by submitting a drawing rather than answering "Does E occur?" The correctness of the agent's answers can be assessed objectively enough in terms of a suitable function that quantifies the degree of match between the randomly generated picture and the agent's drawing, and the value of that function serves as the answer. The experiment can be run while varying the deadline and the kind of agents. My prediction is that for competent human agents, there will be a marked asymmetry such that the agents whose drawings must be submitted with deadlines before E will match the picture much worse than agents whose drawings can be submitted a couple of hours after E. Furthermore, the success rate of the early-deadline human agents will not be surpassed significantly by any other kind of early-deadline agent, including human agents equipped with all the latest technology. Because current technology includes no time machines or Galadriel's mirror, the early-deadline agents will do no better than a guess.

Unfortunately, this modification of the basic epistemic asymmetry experiment is not well designed to distinguish precognizers from random guessers because it is primarily geared towards flagging a difference between an agent before and an agent after the target event. If there are precognizers who are better than random guessers but worse than late-deadline agents who can simply stand in front of the picture and draw a copy, they might be recognized in the experiment by having a higher rate of success than humans who guess wildly. But even a reliably better-than-guessing predictor is not necessarily a precognizer because such differences could also occur if a select group of humans are better guessers merely by knowing the distribution of nouns in the dictionary. What we need instead is a measure of whether some class of agents has some sort of epistemic access to the future that normal people only have towards the past, a measure provided by the prediction experiment.

If some class of agents reliably win money in realistic versions of the prediction experiment, that would constitute prima facie evidence for precognition. However, some imperfections in the experiment are worth noting.

First, all the previously noted respects in which real designated predictors are inferior to ideal predictors apply as reasons to think that the agent succeeded only because of limitations of the designated predictor, not because the agent is surpassing what is ideally inferable about the future from the present state plus the laws.

Second, neither the ideal designated predictor nor its nearly ideal cousins incorporates explicit information about how the pictures are generated. If the randomizing device uses a dictionary to select depictable nouns, an ordinary human could use statistical information about the distribution of nouns in dictionaries in order to make improved guesses, but this kind of information will likely be difficult for a designated predictor to retain implicitly in its representation of the deadline state because of the way correlations among disparate atoms tend to be quickly lost in simulations of chaotic dynamical systems. A related point is that ordinary humans might be able to exploit imperfections in the matching function by drawing images that tend to match a wide range of random pictures to some positive degree whereas the ideal agent merely makes its guess based on what it thinks the probability of the future picture will be. Both of these limitations can be mitigated by using a more appropriate device to generate the pictures. If the pictures are created by filling in an eight by eight square grid where each cell has a fifty percent chance of being black and a fifty percent chance of being white, then it will be easier for the designated predictor to overcome practical limitations in its dynamical calculations, and if we require a perfect picture match, there will be no matching function for the ordinary humans to exploit.

Third, the precognition experiment is not sensitive enough to detect precognizers who see future events through a suitable form of past-directed causation but not well enough to win money against a sufficiently sophisticated designated predictor. Such precognizers would be accomplishing something remarkable because their beliefs about the future would be more accurate than what is warranted on the basis of any psychologically plausible causal inferences that could be made from the agent's knowledge of the past and present. I suspect, though, that once one chooses to evaluate precognition relative to the inferential resources of the agent rather than an absolute standard like a designated predictor, then all sorts of false positives will appear. There would be no principled way to distinguish bona fide precognizers from agents who have internalized some predictive rules of thumb that are reliable in virtue of ordinary causal regularities but where the agent is wholly unaware of the causal principles and does not engage in any relevant causal reasoning. For example, a fortune teller might have hunches based merely on an initial visual impression of her client that result in vague predictions like, "An unexpected obstacle will soon appear in your life, but you will overcome it with persistence." Without the fortune teller having any precognitive abilities, many such predictions are reasonably accurate in virtue of their general applicability to most people. Yet, because the fortune teller is reasoning merely from a hunch rather than from a causal inference, any measure of precognition that measured predictive success relative to the fortune teller's own causal reasoning capacity would be susceptible to mischaracterizing the success as precognition. It is doubtful any agent-relative but external criteria of success can systematically distinguish between (1) genuine precognizers who use precognition to overcome practical limitations in their ability to reason using causal laws and (2) ordinary people whose causal reasoning is so compartmentalized and unsystematic that it is outperformed by instinctual hunches. In order to verify an experimental result as genuine precognition in such a case, one would need to sort through the implementation details to find evidence for some special sort of fundamental nomic link going from the foreseen event back in time to the precognizer (besides the ordinary past-directed nomic links that exist, for example, when the laws are deterministic in both directions of time).

Even though the improved version of the precognition experiment does not adequately distinguish precognition from prediction, it should be clear that because the goal is to measure the ability of agents to exceed what is inferable about the future on the basis of prediction via the laws, the precognition experiment can be thought of as a realistic variant of the prediction experiment, perhaps with the designated predictor having significantly less than ideal predictive capacity. So far as I can tell, running the precognition experiment would not reveal any additional empirical content.

The precognition experiment does, however, illustrate something important about the role of causation in precognition. If we adopt the theory that causation exists in virtue of fundamental laws that govern the continuous temporal evolution of nature and we accept the plausible hypothesis that these fundamental laws are only able to evolve a fully detailed fundamental state while remaining entirely silent about incompletely specified states or events, then we can draw some important conclusions about precognition.

Specifically, we can be fairly sure that there are only two ways for precognition to occur. The first is for there to be some fundamental nomic route from the future to the past that skips over the intervening stages. Although these nomic routes are plausible in models of general relativity where there are wormholes that general allow dynamical development back in time, if we are in a region where there are no such conduits for backwards causation, it becomes strikingly implausible that there is a special fundamental law that backwardly connects the arrangement of the atoms in a future picture to the present arrangement of the atoms in a human's brain. Perhaps if some form of interactive mind-body dualism were true, this connection would be slightly more plausible, but it is difficult to take seriously the hypothesis that there is a special sort of nomic link that hops over previous states in order to impose an influence on certain past mental conditions.

But if there is no special nomic connection to the past, any causation going from the picture back to the alleged precognizer's drawing is causation that exists in virtue of the past-directed dynamical development of ordinary matter and fields, for example, under deterministic laws that are time-reversible. Because of the way the experiment is set up, success would require a lawful development that establishes a reliable correlation in the past between the choice of picture made by the randomizing device and the mental state of the agent. But this conflicts with future-typicality because such a correlation would mean that the superficially random device has a conspiratorial outcome that matches the agent's drawing with the later picture. Thus, if evidence coming from the precognition experiment continually mounts year after year, then the only remaining explanation is that the actual world exhibits a violation of future-typicality that remarkably correlates the precognizer's mental state to the future state to make it superficially like a perception of the future. One might want to say that such cases are not genuine instances of perception because they lack the kind of causation necessary for ordinary perception, but such remarkable coincidences would count as precognition in practice because it would be a reliable pattern despite holding primarily in virtue of fantastic conspiracies hidden in the precise arrangement of subatomic particles.

Because these two implausible scenarios are the only ways for precognition to manifest itself in worlds with realistic laws, I would say it is doubtful that precognition exists.

References

[1] Albert, D. (2000). *Time and Chance*. Cambridge: Harvard University Press.

[2] Dummett, M. (1964). "Bringing About the Past," *The Philosophical Review*, 73 (3), 338–59.

[3] Kutach, D. (2010). "Empirical Analyses of Causation," in *New Waves in Metaphysics*. Hampshire, UK: Palgrave Macmillan, 136–55.

[4] Kutach, D. (2013). *Causation and Its Basis in Fundamental Physics*. New York: Oxford University Press.

[5] Mackie, J. L. (1973). *The Cement of the Universe*. Oxford: Oxford University Press.

[6] Scriven, M. (1956). "Randomness and the Causal Order," *Analysis* 16 (5), 104–10.