An Opinionated Introduction to the Philosophical Foundations of Bayesianism

Kenny Easwaran

Chapter 4: Numerical confidence

Chapter 1 of this book made the case that there is a useful role in epistemology for an attitude of confidence that can be higher or lower. Chapters 2 and 3 considered how we might understand this attitude's role either in governing action or in representing the world truthfully. Both of these roles seem to lead to the idea that degrees of confidence come on a numerical scale satisfying the probability axioms. That is, for any propositions p and q:

- $C(p) \geq 0$.
- If the agent is certain of p, then $C(p) = 1$.
- If the agent is certain that p and q aren't both true, then $C(p \vee q) = C(p) + C(q)$.

However, many philosophers have claimed that this sort of probabilism requires unrealistic precision and numbers in the head, and therefore it can't be correct either as a theory of actual people, or even of the sorts of models we might aspire to emulate. In this chapter I will consider several of these objections in order to defend the idea of degrees of confidence as precise, numerical probabilities. Although I don't have positive evidence from empirical psychology to suggest that actual humans really do behave this way, I will argue that the sorts of features philosophers usually point to in arguing against precise numerical probabilism are actually quite compatible with it. Whether or not we in fact have this sort of confidence, creatures that do have it wouldn't have to be notably different from us.

This discussion will also be useful in clearing up some misconceptions about what it could mean for degrees of confidence to be precise numerical probabilities. I will use many tools from the theory of measurement in social science and the philosophy of science, to compare numerical degrees of confidence to many other features of the world that we feel happy representing numerically. Probabilism itself will be seen to be something like the Fahrenheit scale for temperature — it can accurately represent the underlying facts in a precise and numerical way, even though many aspects of it are pure conventions that could have been represented otherwise.

There are three main families of objections to precise numerical probabilism that I will address in this chapter. The first set of objections focuses on the structure of the arguments from the previous chapters in favor of probabilism, claiming that these arguments couldn't tell us about confidence itself, but only about how we choose to represent it. The second set of objections claims that confidence (or any related notion of "partial belief") could at most explain a small amount of the psychological behaviors that are claimed for it. The third set of objections claim that in any case, any numerical measurement of confidence could at best be extremely imprecise, so that using precise real numbers is drastically misleading.

1 Representationalism

The first family of objections is exemplified by Zynda (2000) and Meacham and Weisberg (2011). They focus on the representation theorem arguments (discussed in Chapter 2), and claim that these can't establish the literal truth of the claim that degrees of confidence either do or should satisfy the axioms of probabilism. At most, they claim, these arguments can serve a sort of heuristic purpose. I will respond by clarifying the role of numbers in the probabilistic representation of degrees of confidence. I will argue that the challenges raised by these authors is no more significant than analogous challenges we can raise to the role of numbers in our thinking about temperature, or distance, or many other physical quantities. While their challenges don't obviously apply to the accuracy arguments discussed in Chapter 3, I think that considering these arguments can help clarify our thinking about the role of numbers there as well, and what we must assume about mental representation in order for that argument to justify probabilism.

Zynda phrases the representation theorem argument for probabilism as using the following three premises:

> The Rationality Condition: The axioms of expected utility theory are the axioms of rational preference.
>
> The Reality Condition: If a person's preferences can be represented with a set of degrees of belief that obey the probability calculus, then the person really *has* degrees of belief that obey the laws of the probability calculus.
>
> The Representation Theorem: A person's preferences satisfy the axioms of expected utility theory if and only if the person's preferences can be represented with a set of degrees of belief that obey the probability calculus.

Together, these premises lead to the conclusion that a rational person really has degrees of belief that obey the laws of the probability calculus. As we saw in Chapter 2, there are significant questions about the Rationality Condition, and whether the particular axioms that it imposes really are requirements of rational preference. One might worry that the correct axioms of rational preference only suffice to prove a weakened form of the Representation Theorem. However, Zynda grants that we might be able to find an appropriate combination of axioms and theorems, and instead focuses on the Reality Condition. Even if we can ensure that a rational person can be *represented* with a set of degrees of confidence that obey the probability calculus, he thinks there may be no sense in which they really *have* such degrees of confidence.

1.1 Four Views

He considers four different views that one might have about degrees of confidence, borrowing terminology from philosophy of science and philosophy of mind. He calls these views "eliminativism", "antirealism", "weak realism", and "strong realism". These four views are a useful framework for understanding what we think of any putative category in the world. The focus of his argument is that a theorist who relies on the representation theorem argument for probabilism must reject strong realism. Meacham and Weisberg go further and say that such a theorist must reject even weak realism. My aim is to defend the idea that strong realism is compatible with all the challenges raised by Zynda, and Meacham and Weisberg.

Eliminativism about a concept is the view that this concept doesn't correspond to anything in the world. This is the view that contemporary scientists take to historical concepts like "phlogiston". Early modern alchemists had developed the historical elemental theory of earth, air, fire, and water, and had come to the conclusion that there were multiple separate elemental earths, and thought of fire as the emission of a substance called "phlogiston" that is part of the composition of various fuels. Through their understanding of the smelting process, they had come to the conclusion that the different metals (copper, iron, silver, gold) were each compounds of a particular earth with phlogiston. They also recognized phlogiston as a substance given off in the respiration of animals. The English scientist Joseph Priestley was able to extract a form of air he thought of as "dephlogisticated", because it promoted combustion and helped the respiration of animals. His work was translated into French by Marie-Anne Paulze Lavoisier, who assisted her husband Antoine-Laurent Lavoisier in various experiments proving that metals actually *gain* mass when converted to the corresponding earth. They argued that Priestley's "dephlogisticated air" was actually a new element, which they called "oxygen", and that combustion and respiration should be understood in terms of absorption of oxygen, rather than emission of phlogiston. In contemporary theory, the different roles played by phlogiston in alchemical theory are played in various places by spare electrons, absence of oxygen, reducing potential, and various other concepts, but phlogiston itself has been rejected. There are similar histories with the "caloric" theory of heat as a fluid (rejected in favor of a statistical mechanical explanation in terms of the energy states of the microscopic constituents of a substance), and the "luminiferous ether" that supposedly conducted light (rejected by Einstein's theory of relativity, in which there is no fixed substance that electromagnetic waves excite).

Eliminativism has been an important view in the philosophy of mind, strongly associated with the work of Paul and Patricia Churchland (Churchland (1981, 1986)), but with important precedents in the work of Wilfrid Sellars (1956) and others. The idea is that just as the concept of phlogiston turned out to be a mistake (albeit a useful one) on the way towards understanding how substances really transform, our ordinary psychological concepts like belief, desire, and perception are also mistakes on the way to a better understanding of the mind.

Zynda thinks that the eliminativist about mental content is unreasonable, because he or she denies the existence of "those features that contribute essentially toward making him or her a person and agent". However, I think thoroughgoing eliminativism about mental concepts generally is a more serious challenge — perhaps personhood and agency can be reinterpreted in the new terms (as combustion was reinterpreted in the oxygen theory), or perhaps we should reject those concepts too once we better understand our nature as physical animals. I think that properly responding to the eliminativist requires a deeper empirical study of the human mind and behavior than I am in a position to carry out. For further discussion of eliminativism, see Ramsey (2013).

In any case, the main thrust of the eliminativist challenge is about actual humans, and as I argued in Chapter 1, I think the project of epistemology is meant to apply not just to actual humans but to other potential cognitive architectures that could count as epistemic as well. In the rest of this chapter I aim to respond to the sorts of challenges raised to the numerical measurement of degrees of confidence by philosophers that don't aim to fully eliminate the concepts of epistemology from the human mind. I claim that the features of human cognition focused on by these philosophers are not unlike what one might expect from Bayesian agents. Thus, unless one wants to reject the application of all ordinary concepts of epistemology to humans, I don't think there is clear reason to reject the application of numerical degrees of confidence satisfying a version of probabilism.

The second view of a concept that Zynda considers is antirealism. Antirealism is sometimes taken as a global view about nearly *all* of the concepts of science. This view is perhaps most fully defended by van Fraassen (1980). On van Fraassen's view, the most our evidence tells us about the world is that our observations proceed *as if* there were subatomic particles and distant galaxies, and *as if* DNA were a molecule obeying particular chemical laws that guided the evolution of life over millions of generations. Since we can't directly observe the microscopic or the extremely distant or the far reaches of biological history, we can't be sure that there is anything real in the world corresponding to these theoretical concepts. At most, we can say that these theories correctly predict the outcomes of experiments.

This sort of global antirealism about theoretical concepts would be no special challenge to probabilism about degrees of confidence. But some scientific theories have involved a more local antirealism. An example is the theory of "quasiparticles" in solid state physics. A substance like copper that is a good electrical conductor can be understood as one in which electrons have a very large number of allowable states at many energy levels, so that they can move very easily between them under the influence of electrical forces. A substance like rubber that is a good electrical insulator can be understood as one in which electrons have very few allowable states, so that they basically can't move at all. A semiconductor like silicon is a substance in which there is a set of low energy states that electrons can easily move between, and also a set of high energy states that electrons can easily move between, but it is difficult to get an electron to move between the low energy and high energy states. Ordinarily, all the low energy states are filled, and all the high energy states are empty, and the negative charges of these low energy electrons are balanced with the positive charges of the atoms. But if you introduce an extra electron into one of the high energy states, then it can flow easily through the silicon as a negatively charged particle, without interference from the sea of electrons in the low energy states.[1] Interestingly, if you instead *remove* an electron from one of the low energy states, this allows nearby electrons to move easily to fill the hole. Instead of thinking about the slight movement of lots of electrons, it becomes easier to think of the *hole* as a "quasiparticle" that moves through the silicon as if it were a positively charged particle. Although only the electrons are *really* there, it's sometimes easier to talk about what's going on by talking about the holes.

Interestingly, this theory of holes initially arose not in the discussion of semiconductors but as the English physicist Paul Dirac's explanation of antimatter. He noted that if we pretend that the vacuum of space has this similar separation between high energy and low energy states for electrons, and pretend that there are electrons filling all the negative energy states, then we could observe holes moving as positive charges just as electrons move as negative charges. If an electron encounters a hole, it can "fall down" to the lower energy state, releasing energy, and destroying both the free electron and the hole. This behavior of holes is exactly the observed behavior of positrons — the anti-particle corresponding to electrons. In semiconductor physics it is common to think of the hole as a fictional entity that simplifies our discussion of the behavior of the real low energy electrons in different types of conduction. In vacuum physics it is common to think of the sea of electrons and the excess low energy states they fill as a fiction that can help our thinking of the real antimatter. Van Fraassen would insist that all ways of thinking about the situation are equally fictional, while a local anti-realist would suggest that one picture is the real picture, and the other is a useful fiction.

The difference between anti-realism and eliminativism is that the eliminativist says that the concept is confused and leads to errors, while the anti-realist allows that although the concept corresponds to nothing real in the world, it

[1] More accurately, there *is* interference from the sea of electrons in low energy states, but the net effect of this interference is that the high energy electrons moves *as if* it were an electron of higher mass in empty space. This is another anti-realist aspect of the treatment of semiconductors.

allows one to make the same predictions and explanations as the correct theory.[2] Anti-realism in this sense is not so common in epistemology and the philosophy of mind, though the term "anti-realism" is sometimes used to mean what Zynda and I call eliminativism. However, in my paper, "Dr. Truthlove, or, How I Learned to Stop Worrying and Love Bayesian Probabilities" (Easwaran, 2015), I showed how a particular realist view of the concept of full belief could underpin an anti-realist story of this sort about Bayesian degrees of confidence satisfying the probability axioms. I didn't give any reason to think this sort of view would be true, but aimed to use it to show that a realist about full belief should not be a total eliminativist about the Bayesian view.

Zynda's main focus however is the distinction between what he calls "strong realism" and "weak realism". While these views agree that the concept involved corresponds to something real, they disagree about how fundamental the concept is. A biological population is a real thing, but it is composed of many individuals that compete and reproduce with each other. The individuals are more fundamental than the biological population, but there are useful laws that can be expressed at both levels. The planets of the solar system are real objects, but they are composed of chunks of solid and gas (and in a few cases liquid) that move around in ways that are more fundamental. As Zynda points out, for some purposes we can treat the center of mass of the Jupiter system as the object that orbits the sun, but while this center of mass is real, its behavior is in some fundamental sense to be explained by the behavior of all the gases and metals and ices of the planet and its moons.

The distinction between strong and weak realism is comparative here. Neither Zynda nor anyone else (except perhaps some sort of Cartesian dualist about the mind) thinks that degrees of confidence are fundamental irreducible entities in the world. Everyone agrees that degrees of confidence, if they really exist, are somehow constituted (at least in humans) by the behavior of bodies and neurons, and their interaction with the world. Like anything else psychological or social, the extent of the physical world that constitutes these states is unclear. There is such a thing as the exchange rate between the US dollar and the Japanese yen, but is it just constituted by the attitudes of central bankers, or does it depend on the behavior of all international commerce between the United States and Japan, or does it even depend in minute ways on the degree to which I am prepared to act on an idle fantasy of purchasing a ticket to hike up Mt. Fuji? Similarly Putnam (1975) argues that the contents of a person's mental states depends not just on the states of her own body, but also on her history of interaction with the objects her thoughts are about, and her tendencies to defer to various social 'experts' on the meanings of her concepts. Furthermore, Clark

[2] Depending on how detailed the predictions and explanations required are, these two views may shade into one another. The treatment of fuels as highly phlogisticated substances (rather than their modern treatment as reducing agents that combine with oxidizers) may be sufficient for some purposes that aren't too sophisticated. Meanwhile, the treatment of a hole as a quasiparticle behaving like a positively charged electron in a semiconductor does run into some difficulties near the edges of the solid, where it might have to interact with electrons from a substance with different energy levels.

and Chalmers (1998) argue that one's mental state may partly be constituted by the various physical and social structures that shape one's mental life, from notepads to smartphones to communities.

But the question for Zynda is whether the preference that the axioms of a representation theory apply to is more or less fundamental than the degrees of confidence that form part of the representation that is proved to exist. The strong realist about degrees of confidence will say that however degrees of confidence are constituted in the world, they are part of what gives rise to the preferences of an agent. The weak realist, on the other hand, will say that preferences are the more fundamental mental state, and that degrees of confidence, and perhaps even belief and desire themselves, are just useful summaries of a more fundamental notion grounded in something like our choice behavior.

I take it that something like the strong realist view is the one most people land on, once they accept the notion of degree of confidence as something real. It certainly seems natural to say that one prefers one act to another *because* one is more confident that it will lead to a good outcome. But Zynda, like many other philosophers and economists that work with representation theorems, wants to argue that this is an illusion. The only way that they say we can make sense of degrees of confidence is as something implicitly defined by our preferences, rather than as something more fundamental that explains our preferences.

1.2 Zynda's Challenge

Zynda argues that the strong realist can't support the Reality Condition of the representation theorem argument. To do this, he describes two agents, called Leonard and Maurice. Leonard and Maurice have all the same preferences over gambles — they both prefer to receive a prize if a fair coin lands heads rather than if a fair die rolls a 3, and both prefer to receive a prize if a fair die comes up less than 5 than if a fair coin lands heads.

However, they report their degrees of confidence quite differently. Leonard reports his on a scale from 0 to 1, and says his confidence that a fair coin lands heads is .5, while his confidence that a fair die rolls a 3 is 1/6. His degrees of confidence satisfy the probability axioms — they are non-negative, he has confidence 1 in things he is certain of, and his confidence in a disjunction is the sum of his confidences in the disjuncts, if he thinks they are incompatible. Maurice, however, reports his degrees of confidence on a scale from 1 to 10, and says his confidence that a fair coin lands heads is 5.5, while his confidence that a fair die rolls a 3 is 2.5. His degrees of confidence satisfy a different set of axioms — they are at least 1, he has confidence 10 in things he is certain of, and his confidence in a disjunction is 1 less than the sum of his confidences in the disjuncts, if he thinks they are incompatible.

Because these agents have the same preferences, they both satisfy the Rationality Condition of the representation theorem argument. Thus, by the Representation Theorem, both can be represented as having degrees of confidence satisfying the probability axioms. In order to save the Reality Condition therefore, we must deny the significance of Maurice's self-reports of violating the probability axioms. A weak realist might say that since degrees of confidence are dependent on preferences, the self-reports don't matter, and the degrees of confidence of both agents really are those reported by Leonard. But a strong realist must say that the degrees of confidence are in some sense prior to the preferences. In order to save the Reality Condition, a strong realist must therefore say that whatever it means to "obey the probability calculus" is something that is shared by Leonard and Maurice. In particular, none of the numerical features of probability are literally true.

Meacham and Weisberg (2011) argue that this sort of view would deny too much of probabilism to be worth saving as a theory of real degrees of confidence. They note that while Leonard and Maurice's self-reports differ just by a simple transformation of adding 1 and multiplying by 9, one could proliferate examples with other mathematical transformations. If someone else had degrees of confidence that were the squares of those of Leonard, then this person would have confidence 1/36 that a fair die rolls a 3 (where Leonard has 1/6), and 1/4 that a fair coin comes up heads (where Leonard has 1/2), and 25/36 that a fair die rolls less than a 6 (where Leonard has 5/6). Both people agree that rolling a 3 is least likely, getting heads is in the middle, and rolling less than a 6 is most likely. But while Leonard says that these three are equally spaced, this other person says that her confidence of getting heads is substantially closer to her confidence of rolling 3 than to her confidence of rolling less than a 6. At most, Meacham and Weisberg say, we can preserve the reality of ordinal comparisons, saying which events are more likely than others, but not saying anything about cardinal comparisons of how *much* more likely one is than another.

Furthermore, Meacham and Weisberg argue that this view leaves us unable to compare the degrees of confidence of different people. If degrees of confidence are just a feature of what each person has, and each can be transformed in a different way, then there is no way to compare a degree of confidence one person has with a degree of confidence of another person.

However, I will argue that by properly understanding the analogy between the measurement of degree of confidence and the measurement of quantities like distance, temperature, and so on, we can recover interpersonal comparisons of confidence. Furthermore, we can recover various notions of how *much* more confident one is of one proposition than another, though I think there are some conceptual worries about the extent to which this notion really is significant. While none of this establishes that degree of confidence is *in fact* prior to preference, it is just meant to show that there are no significant problems on this front from assuming that it is.

1.3 Measurement theory

The aim of this section is to argue that degree of confidence is much like temperature. Donald says that water freezes at 32 degrees, and that room temperature is 68 degrees, while Justin says that water freezes at 0 degrees, and that room temperature is 20 degrees. We take this to be a difference in reporting, but not a difference in the strongly real temperature underlying the two descriptions.

As I will argue towards the end, there is actually a similar ambiguity to the notion of "how much hotter" as there is in "how much more confident". But before I get there, I will use distance as an example to show how the numerical representation of real parts of the world works.

The central ideas were set out most systematically by Krantz et al. (1971), and my presentation follows much of their work. The concepts of length, mass, temperature, and so on aren't literally numerical. Rather, in each case, there are some real physical relations that happen to obey various laws, and since those laws have the same structure as some arithmetical laws, we can choose to represent those physical relations with numbers. Some features of this representation correspond to the real physical laws, while others are just conventional choices made to simplify our calculations. The same is true for degree of confidence, and the challenge is to figure out which aspects of the representation are to be taken literally and which are to be taken as conventions. I will outline the situation for distance, then discuss some additional features that show up in the discussion of temperature, and then draw analogies for degree of confidence.

1.3.1 Distance

For distance, we assume there are two fundamental, real relations. The first is a comparative notion that lets us compare pairs of locations. If A and B are two locations, and C and D are two locations, we imagine rigid rods stretching between those locations, and think about what would happen if we were to line them up parallel to each other, starting from the same point. If the rod that started stretching from A to B would stick out past the end of the rod that started stretching from C to D, then we say that $AB \succ CD$. Some amount of fictionalization and idealization is needed for this — there is no rod stretching from Houston to Dallas, and there is no rod stretching from Los Angeles to New York, and even if there were, no one could move them to line them up. But what is important is that there is *some* feature of the world that means that if these rods *did* exist, and someone *could* move them, the one from Los Angeles to New York would stick out past the one from Houston to Dallas. Maybe you'd need to somehow move the Rocky Mountains, and relocate the population of the central United States to move these rods without breaking them. But the claim is that something like this is part of what we really mean when we say that the distance from Los Angeles to New York is 2470 miles while the distance from Houston to Dallas is only 239 miles.

The numerical representations of these distances need to correspond to this real relation. We will want to represent each pair of locations with a number in a way that $AB \succ CD$ iff the number corresponding to AB is greater than the number corresponding with CD. We will also eventually impose some other constraints on this representation, but in order for this sort of representation to work at all, we need the world to govern the behavior of these hypothetical rigid rods in particular ways. The first thing that is needed is that since no number is greater than itself, it needs to be the case that if you start with two copies of a rigid rod, there is no way to move them so that one of them sticks out past the other. This is basically what we mean by "rigid", but it is important that there really be some feature of space itself (or at least, the part of space we are considering) that supports this feature. If there are strange wormholes or ways of moving objects relativistically that deform them in this way, then our notion of distance will break down if we try to apply it in those contexts. The second thing that is needed is that if $AB \succ CD$ and $CD \succ EF$, then $AB \succ EF$. That is, if there were three of these rigid rods, and the first could be moved to stick out past the second, and the second could be moved to stick out past the third, then the first could be moved to stick out past the third. In combination with the first rule this means that no matter how you move two rods, whichever one sticks out past the end of the other always will, no matter how you line them up. This seems quite natural, but it is an empirical feature of space that guarantees this, and not mathematics itself. One other important thing to note — since we are imagining such hypothetical rods connecting *any* pair of points in space, there is also a hypothetical "rod" AA from a point to itself, which we might think of as empty. It is important that for any two distinct points A and B, AA is not longer than BB, but AB is longer than either AA or BB.

The second real feature of the world that is needed to fix our numerical representation of distance is the idea of what happens if two of these hypothetical rigid rods are laid end-to-end in a straight line. The resulting "concatenation" of rods should behave just like one of these rods, and can thus be compared to individual rods. If AB and CD are two such rods, then we will represent this concatenation by $AB \circ CD$. Again, these rods don't really exist, and it's even harder to imagine lining a pair of them up in a row and making the result rigid than it is to imagine just comparing two of them. But I claim that something about this idea is that we mean when we note that the distance from Houston to Austin of 152 miles, added to the distance from Dallas to San Antonio of 248 miles, is less than the distance from New York to Chicago of 737 miles.

We want our numerical representation to be such that $AB \circ CD$ is represented by the number that is the sum of the numbers representing AB and CD. In order for this to work, there are several more features of these rods that the world needs to ensure. First, since $x + y = y + x$, it needs to be the case that sticking two rods together in one order, and sticking them together in the other order, give rods of the same length. Second, it needs to be the case that if $AB \succ CD$ and $EF \succ GH$, then $AB \circ EF \succ CD \circ GH$. That is, there is no way to put together two shorter rods and get something longer than what you get if you put together two longer rods. Although these features of the physical world are so fundamental to our understanding of the world that they *feel* like mathematical truths, they are in fact substantive assumptions. In fact, much of our intuitive understanding of arithmetic probably derives from our physical experience with distances and other real-world concepts that behave in these analogous ways.

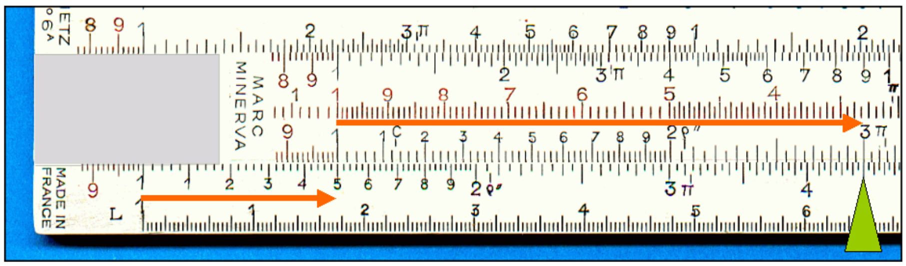
Once we have all of these features, we can begin to assign numbers to distances. For the trivial "rod" AA, we see that sticking it to the end of any other rod doesn't change the length. So if concatenation is represented by addition, then this trivial rod must be represented by the number 0. If we choose some other specific rod and call it 1, then we can begin to approximate other distances by seeing how many times this rod can go into it. Something like this is behind the idea of using one's foot to measure distances in feet. But the assumptions we have made so far don't quite guarantee that we can get precise numbers to correspond to all distances.

One quickly notes that whatever rod one uses, there are distances that don't exactly line up with any number of copies of that rod. If you lie down on the ground and mark the length from your head to your toes, it is quite unlikely that you can fit exactly 5 feet or exactly 6 feet in that length. Most likely there will be some space left over. We are used to dealing with this by choosing some shorter rod (perhaps the first joint of one's thumb, which one can call an "inch"), using it to measure the leftover, and seeing how many times this rod goes into the first one. To be sure that this technique can always work to get more and more precise numerical approximations of distance, we need it to be the case that for every rod, there is another rod that goes into it at least twice. That is, for every AB, there is some CD with $AB \succ CD \circ CD$. This question of whether space is actually arbitrarily divisible in this way has been a philosophical question since antiquity, and lies at the heart of paradoxes from Zeno to Kant. But for most purposes, we can at least treat distance as if this sort of divisibility obtains.

The other fundamental assumption that is needed for this method to work is that for any measuring rod we choose (other than the trivial rod AA), and for any distance we want to measure, it has to be possible to exhaust this distance by some finite number of copies of the measuring rod. No distance can actually be infinitely far, and no non-empty rod can actually be infinitely short. This is known as the Archimedean principle of distance, and again is part of our basic understanding of the world derived from experience. If you are standing in Los Angeles, it can *feel* like putting one foot in front of the other gets you no closer to New York. But if three feet are as long as a yardstick, and 22 yardsticks are as long as a surveyor's chain, and 80 chains make a mile, and there are 2470 miles from Los Angeles to New York, then there is some number of feet that gets to New York. There is even some number of millimeters to the edge of the galaxy!

Assuming that distance actually has all of these features, so that moving rods around doesn't change their ordering, and so that rods can be concatenated in appropriate ways, and every distance can be subdivided by every other, we can represent distances with numbers. If we choose any unit to represent with the number 1, then all other distances can be represented in a unique way if we want concatenation to be represented by addition, and greater distance by a greater number. The choice of unit is arbitrary, and once that choice is made, the rest of the representation is fixed. There is no sense in which the number 2470 is itself part of the distance from New York to Los Angeles — the number is a pure convention, and someone measuring in kilometers would represent the same distance as 3975.

While most modern people are at least implicitly familiar with this much of the conventionality of the numerical representation of distance, there is a bit more that can be abstracted away from the numbers. We made an arbitrary choice to represent concatenation of rods by *addition* of numbers, rather than some other operation. A slide rule makes use of this arbitrariness to help us calculate multiplication or other functions. (Image CC by Jean-Jacques MILAN at Wikimedia)



Where the ordinary rule at the bottom of this slide rule puts the numbers 0, 3, 6, etc. at equal intervals, the slide rule just above it puts the numbers 1, 2, 4, etc. at equal intervals, with the other numbers (and fractions of them) at appropriate places in between. With an ordinary ruler, sliding a length indicated by 4.9 to the end of a length indicated by 1.8 will yield a combined length indicated by their sum of 6.7. But with a slide rule, sliding a length indicated by 3 to the end of a length indicated by 1.5 will yield a combined length indicated by their *product* of 4.5. Sophisticated slide rules sometimes have multiple types of indication of distances, so that other operations than addition or multiplication can also be indicated by concatenation.

What is essential to the types of mathematical operations that can be indicated by concatenation is that they obey the properties of the actual physical concatenation operation. Concatenating a pair of lengths in either order yields the same total length, so the operation we represent concatenation with must also obey this "commutative" law. Concatenating a pair of longer distances yields a longer distance than concatenating a pair of shorter distances, so the operation we represent concatenation with must be increasing in both of its inputs. Concatenating any distance with the trivial distance from a point to itself must yield the original distance, so there must be some number that serves this null purpose in the operation we represent concatenation with. For addition this number is 0, but for multiplication this number is 1.

If we want to make a slide rule where concatenation represents multiplication, we must mark the beginning of the ruler with 1, and then choose a particular distance, perhaps an inch, to mark with any greater number, such as 10. Points at every inch along the ruler are then marked 100, 1000, etc. To mark other numbers, we need to see how many times they multiply into any of these distances. Since $2^3 = 2 \cdot 2 \cdot 2 = 8$ and $2^4 = 2 \cdot 2 \cdot 2 \cdot 2 = 16$, we see that the number 2 must be marked at a distance where three copies is less than the distance marked 10, but 4 copies is more, so it must be somewhere between 1/4 and 1/3 of an inch. Since $2^6 = 64$ and $2^7 = 128$, we see that 6 copies of this distance must be less than 2 inches, and 7 copies must be more, so it must be somewhere between 2/7 and 1/3 of an inch. By comparing further powers of 2 with other powers of 10, we can get better approximations of where exactly 2 should be marked on this slide rule, and similarly for each other number.

If we had grown up with only slide rules for measuring distance, we might naturally think of distance concatenation in terms of multiplication. This might have only been convenient if our system for writing numbers made multiplication as easy to calculate as addition actually is. Such a number system would require us to deal with astronomical numbers when indicating the distance from New York to Los Angeles, or perhaps we would have had a greater variety of units than just inches, feet, and miles (or meters and kilometers). But had we represented distances in this way, the actual fact that we discuss by saying that New York to Chicago is about three times as far as Dallas to San Antonio would instead by discussed by saying that New York to Chicago is about the *third power* of Dallas to San Antonio.

1.3.2 Temperature

The situation for the representation of temperature has some similarities and some differences to the situation for the representation of distance. We are familiar with the idea that the degree Fahrenheit and the degree Celsius are two arbitrary choices of unit on the temperature scale, just as the mile and the kilometer are for distance. But with distances, there is a special distance (namely that from a point to itself) that we represent with the number 0 on every scale, as long as we measure distance additively. (On a slide rule, that distance is marked with the number 1.) Our familiar everyday temperature scales instead make different conventional choices of what temperature is marked as 0.

Just as comparison of rods, and concatenation of rods, are the real (or at least, potential, in the case of long rods no one will ever make) features of the world that govern the numerical representation of distances, there are some real features of the world that govern the numerical representation of temperature. For any two physical systems at equilibrium, if these two systems were to be put in contact with each other, energy would flow from one system to the other, but not the other direction.[3] If energy would move from A to B, then we say that $A \succ B$. As a matter of empirical fact, two copies of the same system would not transfer energy in either direction, and if energy would transfer from A to B and would transfer from B to C, then if A and C were in contact, then energy would transfer from A to C.[4]

The feature of the world that corresponds to the unit of temperature is a bit less familiar, but is approximately summarized by Newton's Law of Cooling. This law states that for any pair of substances (whether water and copper, cast

[3] From a statistical mechanical perspective, this is because the total energy of the two systems is conserved, there are only a finite number of ways for energy to be distributed among all the microscopic units of the two systems, and most configurations of the microscopic units with a given total energy add up to a particular distribution of energy between the two systems, which is generally not the distribution the two systems originally had.

[4] These features of the physical world basically correspond to the idea that systems always tend towards an equilibrium, rather than moving in a perpetual cycle.

iron and onions, or a human body and the air), the rate at which energy flows from the warmer to the cooler is proportional to the temperature difference. A human body at 102 degrees in a room that is 70 degrees will chill just as quickly as a human body at 98 degrees in a room that is 66 degrees. A copper pot at 200 degrees will heat 70 degree water at the same rate that a copper pot at 300 degrees will heat 170 degree water. Coffee at 130 degrees in a 70 degree room will cool off twice as fast as coffee at 100 degrees in a 70 degree room.[5] If one can measure the rate of transfer of energy between a pair of substances, then one can use this to mark off temperatures separated by the same distance according to this law. Whatever reference pair was used to define this difference then gets marked as a degree. If one is just measuring heating and cooling, the choice of zero is arbitrary, and the size of the degree sets the scale of energy flow for any given pair of substances.

However, temperature governs not just heat flow, but also the expansion of gases, and the efficiency of engines that turn heat into work or vice versa (whether through the use of expanding steam, exploding gasoline, expansion and compression of freon, or conversion of solar radiation into electricity). Through the work of the French engineer Sadi Carnot, and later work of William Thomson (granted the title of "Lord Kelvin", by Queen Victoria), it became clear that the expansion of a gas, and the efficiency of an engine or heat pump, depend not on the *difference* of two temperatures, but rather on their *ratio*, when expressed in a scale with a suitable definition of 0 that we now call "absolute zero", at about -273 degrees C or -460 degrees F. As a result, there are *two* notions of the relevant temperature difference. For purposes of figuring out how quickly energy transfers between objects, -270 C is closer to 0 C than 1000 C is — energy flows nearly four times as quickly from a source at 1000 C as it does to a sink at -270 C. But for purposes of figuring out how efficiently an engine runs, 0 C is over a fifth of the way to 1000 C, while -270 C is only about a thirtieth of the way to 0 C. Cooling the exhaust vent/intake valve of an engine can improve efficiency more than increasing the temperature of the heat source.

For some purposes in thermodynamics, it turns out that it's more useful not to represent temperature even on the Kelvin scale with absolute zero, but in terms of its *reciprocal*, sometimes called "coldness". That is, while the melting point of ordinary ice is at a temperature of about 273 Kelvin, we can say that the coldness of this state is about .00366. Liquid nitrogen has a temperature of about 77 Kelvin, and thus a coldness of about .013. Liquid helium reaches a temperature of about 4 Kelvin, and thus a coldness of about .25. When describing progress in low temperature physics, it might seem surprising to hear just how much harder it is to get from 4 K to 1 K and then to .01 K and .000001

[5] A complexity of temperature difference that doesn't exist for distance is the dependence on the substance. The rate of energy transfer from water is much faster than that from air, which is why (all temperatures in Fahrenheit) an 80 degree bath can feel cooler than an 80 degree room (because you're losing energy to it faster) and a 120 degree hot tub can be deadly while a 120 degree dry sauna is barely getting started (because water transfers energy to you so much faster than air). With metal it's even faster, which is why metal so often feels more extremely cold or hot to the touch than the water it contains at the same temperature.

K. But in terms of coldness, this is the progress from .25 to 1 to 100 to 1,000,000. This representation helps make it clear why the third law of thermodynamics says it is impossible to bring a system to absolute zero — this would involve bringing its coldness to infinity![6]

The fact that temperature and coldness are equally good numerical representations of the same subject matter, and the fact that temperature differences and temperature ratios both enter into important physical laws, are the two important analogies I want for degree of confidence. There is no disputing that temperature is in some sense a real feature of physical systems.[7] This feature is appropriately described by fairly precise real numbers. However, the choice of Fahrenheit or Celsius is an arbitrary convention. For some purposes, we can do better by using a scale with absolute zero, like the Kelvin scale. But for those purposes, there is still an arbitrary convention of using temperature rather than coldness to describe the system. The information content is the same, though they emphasize different sorts of difference.

1.3.3 Confidence

Degree of confidence has many of these same features. So far, when I have represented degrees of confidence numerically, it has been with numbers that satisfy the probability axioms. As discussed in Chapter 2, these are the numbers that emerge when one describes the bet a person would consider fair in terms of the price paid for a ticket that pays \$1 if the relevant proposition is true. If a race has three horses, A, B, and C, and one is equally confident that each would win, then one shouldn't be willing to pay more than about 33 cents for a bet that pays a dollar if the given horse wins (because otherwise one would be willing to buy tickets on all horses for more than the possible winning) and one shouldn't be willing to sell such a bet for less than about 33 cents (because otherwise one would be willing to sell tickets on all three horses for less than the amount one would foreseeably have to pay out).

However, if one actually goes to a racetrack to place bets on a horse, one will usually find these prices listed quite differently. I am told that at racetracks in continental Europe, one will find bets listed in terms of the payout for a

[6]From statistical mechanics, it becomes clear that in a system of interacting particles with a given total amount of energy, there is an equilibrium distribution of what fraction of the particles will have a given energy level. It can be calculated that there will be some constant β such that the number of particles at energy level E will be proportional to $e^{-E\beta}$, the Boltzmann distribution. If β is larger, then many more particles are in low energy states than high energy states. As β gets smaller, more particles enter higher energy states. When separate systems individually in equilibrium are brought together, energy will flow from the system with lower β to the system with higher β, which shows that β has the opposite ordering of the characteristic comparison of temperature. It is an empirical fact that β times temperature equals a constant, which is known as Boltzmann's constant. Thus, β is a representation of the coldness, which falls more directly out of statistical mechanics than temperature itself. For more details, see Nash (1974).

[7]There are some limitations in that temperature only strictly applies to systems that are internally in equilibrium, and that systems need to be large enough to have statistical distributions in order to be characterized by temperature. But ordinary macroscopic systems usually have these features.

ticket that costs 1 Euro to buy. If the probability of horse A winning is 1/6, the probability of horse B winning is 1/12, and the probability of horse C winning is 3/4, then a European racetrack would represent the odds as 6.00, 12.00, and 1.33 respectively. However, at a British racetrack, the bets will be listed with odds indicating the ratio of the respective risk of the bookie and the bettor — if a ticket can pay 6.00 on a price of 1.00, then the bookie risks losing 5 units while the bettor risks losing 1 unit. Thus, at a British racetrack, the odds would be represented as 5 : 1, 11 : 1, and 1 : 3 respectively. American racetracks represent these even more strangely. Bets on events with probability less than 1/2 are represented with positive numbers, while those with probability greater than 1/2 are represented with negative numbers. Positive numbers represent the possible profit on a bet that costs \$100, while negative numbers represent the cost of a bet that yields a possible profit of \$100. Thus, for the same three horses at the same probabilities, an American racetrack would list the odds at +500, +1100, and −300 respectively, since a bet of \$100 on horses A or B could yield winnings of \$600 or \$1200 (and thus profit of \$500 or \$1100), and a bet of \$300 on horse C could yield winnings of \$400 (and thus profit of \$100).

These different odds listed at different racetracks form a real-life Leonard and Maurice situation. They are representing the same underlying degrees of confidence on different numerical scales. The significant question is what features the underlying degrees of confidence have that are being equally well represented by the three numerical scales. Distance must be stable under moving rigid rods, in a way that preserves various features of comparison and concatenation, in order to be properly represented by our ordinary additive scale, but then it is also equally well represented by the multiplicative scale on a slide rule. Temperature must be related in systematic ways to the transfer of energy and efficiency of engines in order to make sense of the ordinary scales, but it can then also be represented by coldness. To address Meacham and Weisberg's concern about interpersonal comparison of confidence, we need to understand this underlying real feature of confidence in a way that lets us represent the confidences of different people on the same scale.

And in fact, a version of this work was already done even before the general work on measurement theory by Krantz et al. (1971). In his (1946), the physicist Richard Cox gave a series of features that confidence must have in order to be represented by probabilities, and he showed that any system that has these features can be represented either by probabilities, or by any monotonic numerical transformation of them like the European, British, and American odds.[8] To fully state Cox's theorem requires the notion of *conditional* degree of confidence,

[8]Technically, Cox's Theorem contains a missing step, but this has been clarified and fixed by the computer scientist Joseph Halpern (1999a,b). Importantly, for the proof to work, for every proposition that a person has a non-minimal degree of confidence in, there must be another one that the person is less confident in while still being non-minimal. The set of possible degrees of belief must be infinite in order to establish unique reference points at all levels of the scale. This is no more problematic for finite beings like us than the fact that our finite vocabularies allow us to express infinitely many sentences in our native language, and that for every event, someone can consider the question of whether that event will happen and *also* an independent coin will come up heads.

which will be the subject of Chapter 6. But for now, it is sufficient to say that Cox required basically just that there be some systematic way that one's confidence in a proposition is related to one's confidence in its negation, and some systematic way that one's confidence in a conjunction $A \wedge B$ is related to one's confidence in A, and one's conditional confidence in B, given A. Given such systematic relations for a specific agent's psychology, there is always a unique way to represent her degrees of confidence on a probability scale, and a unique way to represent her degrees of confidence on any other numerical scale that one chooses to use. If we can justify Cox's assertion of such systematic relations, then interpersonal comparisons are not a problem. Many physicists, like Cox and Jaynes (2003), have just assumed that Cox's assertions are "common sense", so there is no need to justify them, but I think there is some important work to be done here.

This is how I see much of the work of the theorems from Chapters 2 and 3. The role of confidence in guiding action, or in accurately representing the world, should not be used to justify the specific use of the probability representation on a scale from 0 to 1. Instead, it should be used to explain the systematic structure of degrees of confidence and show that it *can* be represented on this scale, though others also work. The representation theorem argument that Zynda, and Meacham and Weisberg, consider does not by itself justify strong realism about degree of confidence. But if we have reasons for thinking that there is such a strongly real psychological state that guides one's decision making, then the representation theorem argument gives us a way to understand its structure that justifies its measurement by probabilities, or by any of the other types of odds scales that have been used in gambling.

If one instead takes closeness to the truth, as in Chapter 3, as the fundamental role for this psychologically real state, then the argument must work differently. The argument attributed there to Joyce (and developed further by Pettigrew and others) requires some very specific numerical features in a measure of accuracy, and argues that if one's degrees of confidence do a good job of being accurate, then they will satisfy the probability axioms. If the argument as phrased by Joyce is correct, and accuracy really is measured in his way, then it would be incorrect to represent degrees of confidence by the odds scales mentioned above. Since these scales are clearly reasonable, Joyce's argument must prove too much. There must be a way to weaken his assumptions, if the general strategy is right.

And in fact there is. Lindley (1982) proved a more general theorem. Although Lindley didn't give the philosophical interpretation that Joyce did, his theorem is more general. While Joyce made six very specific assumptions about how accuracy should be measured, Lindley makes just a couple assumptions, that are more general. The details aren't essential to the argument here, but there are two important features. One is that with his weaker assumptions, Lindley is able to show not that degrees of confidence *do* satisfy the probability axioms, but merely the claim proven by Cox that their scale can be *converted* in a unique way to the probability scale. The second feature is that Lindley needs a numerical scale for measuring accuracy (with assumptions about additivity and differentiability), but gives no indication of what the numbers mean in real terms. The considerations of this chapter show that the accuracy framework needs better explanation of what the numbers really mean in order to be considered a complete argument for probabilism.

There is another way to establish a numerical scale for confidence by considering Meacham and Weisberg's second question of how *much* more confident one is of one proposition than another. They claim that the sort of flexible numerical representations considered here (where each form of odds is equally good as the representation as probabilities) undermines meaningful ideas of comparing not just confidences in propositions, but *differences* in confidence. I think this idea of comparing differences in confidence is perhaps more subtle than they suggest. But I think that one way of considering these differences leads to another natural use of measurement theory to produce numerical structure for degrees of confidence.

1.3.4 Differences in confidence

As discussed by Fitelson (1999), there are actually a variety of different measures of difference in confidence that have been proposed by Bayesians, even within the probability representation by itself. If degree of confidence x is greater than degree of confidence y, and degree of confidence y is greater than degree of confidence z, then everyone agrees that the change from x to z is a bigger change than the change from x to y or the change from y to z. The only case mentioned by Meacham and Weisberg (Hosiasson-Lindenbaum (1940)'s influential Bayesian treatment of the ravens paradox) happens to be a case of this form. But it is harder to compare differences in confidence when the differences involve two completely separate pairs of degrees of confidence.

For instance, in one study, 34% of 50 year old American smokers eventually developed cardiovascular disease, while 27% of nonsmokers did. (Lloyd-Jones et al., 2006) In another study, 5.2% of male Polish former smokers died of lung cancer, while 1% of Polish men who had never smoked did. (Brennan et al., 2006)[9] To simplify calculations, I'll treat the change in cardiovascular disease as 33.3% to 25%, and the change in lung cancer from 5% to 1%. In terms of probability, the change in cardiovascular disease was about 8%, while the change in lung cancer was only 4%, so you might say that the particular observations yielded twice as large an effect for smoking as for lung cancer. However, smoking brought the *odds* of getting cardiovascular disease from 3: 1 to 2: 1, it brought the odds of lung cancer from 99: 1 to 19: 1. It is also common to report changes in terms of ratios rather than differences. The ratio of probabilities (reported in medical journals as a "risk ratio") in the cardiovascular disease study is 4/3, while the ratio of probabilities in the lung cancer study is 5/1. And the odds ratio in the cardiovascular disease study is 3/2 while the odds ratio in the lung

[9] Numbers were much higher for people who continued smoking throughout their lives, ranging from 17% in Russia to 82% in Slovakia. The rates of lung cancer in non-smokers in these Eastern European countries were apparently also higher than the rate of 0.2% often observed in the UK or US. I chose the examples I did for convenient numerical values.

cancer study is about $99/19 \approx 5.2$. On either the risk ratio or odds ratio measure, lung cancer was much more affected by smoking than heart disease.

It's not clear to me that there is a univocal answer to which difference should be treated as larger. So I think we shouldn't give in to Meacham and Weisberg's demand to produce a clear distinction between "greater" confidence and "much greater" confidence. Some Bayesians (Christensen, 1999, Vassend, 2018) have suggested that these different measures should be used to compare changes in different contexts. The difference in probabilities may be the more relevant numerical measure of change for purposes involving action plans, while other measures might be better for other purposes. However, as I'll show, there's at least some sense in which the odds ratio (or perhaps better, the logarithm of the odds ratio) is a natural measure for the amount of evidence one has. This may also give another way to say what is real in degree of confidence to get a quantitative treatment started through measurement theory separate from either the decision theory or accuracy frameworks mentioned so far.

Consider a situation in which one is investigating a trick coin from a magic shop. One knows in advance that some coins from the shop are biased with a 2/3 chance of landing heads, and the rest are biased with a 2/3 chance of landing tails.[10] Whatever initial degree of confidence c one has for the coin being biased towards heads, the probability of getting heads on all of the first n flips is $c \cdot (2/3)^n + (1-c) \cdot (1/3)^n$. While the posterior probability for the coin being biased towards heads is somewhat awkward to calculate, the odds ratio is quite straightforwardly 2^n times what it started with. Each tails flip one observes divides the odds ratio by the same amount. Thus, the logarithm of the odds ratio counts the number of observations of independent coin flips one would need to become equally confident in the coin being the other way from how it is now.

Thus, we can imagine setting up a quantitative notion of measurement for *amount of evidence* just as we did for distance initially. Instead of comparing distances between pairs of points AB and CD, we can compare distance between propositions (P, Q) and (R, S) in this evidential sense. For distance, we have to imagine a pair of hypothetical rigid rods stretched between AB and CD, and imagine sliding them towards each other to compare which is longer. For evidence, we have to imagine a pair of hypotheses like the ones about biased coins, and figure out whether the number of independent observations needed to bring it from the confidence one has in P to that of Q would be more or less than the number of independent observations needed to bring it from the confidence one has in R to that of S.

Concatenation of evidence is simple in the case of pieces of evidence that are independent conditional on each hypothesis. When pieces of evidence aren't independent, we have to imagine replacing them with ones that are independent, just as when pairs of points aren't collinear, we have to imagine straightening out the rods so that they are lined up. If these evidential analogs of the movements

[10]Though see Gelman and Nolan (2002) for an argument that such trick coins are actually impossible!

of rigid rods can be made sense of, then we can directly choose some unit of evidence and get a ratio scale for it in the same way that Krantz et al. (1971) gets a scale for distance. This could be another route to measurement theory for degrees of confidence if one can make sense of the relevant notion of comparative evidence without already having the mathematics of probability theory. In fact, something like this discussion was given by I.J. Good and Alan Turing in their early wartime work on codebreaking, though it was't published until much later. (Good, 1985) Interestingly, Vassend (2015) gives a very different argument that the logarithm of the change in the odds ratio is the only confirmation measure that behaves well from a measurement theory perspective.

To summarize, I think it's not clear what the best measurement theoretic foundations are for degree of confidence. However, several good starts are available. In any case, Zynda's worry about alternative numerical representations is not a worry for the strong realism of degree of confidence — it is a universal feature of all numerical quantities, whether psychological or physical. Meacham and Weisberg's worries about these numerical representations can be addressed if good foundations for the measurement of confidence are available. Each person's degrees of confidence will have a unique representation as a probability function (and also a unique representation as British odds, or European odds, or American odds), and this representation will suffice for interpersonal comparisons of confidence. Comparison of differences of confidence works fine in the one case that is most pressing (that of saying when one piece of evidence is stronger than another for the same proposition), while in other cases there may not be a univocal notion. But there may be a measurement theoretic foundation for a notion of confirmation that uniformly answers that challenge as well, based on the odds ratio.

2 Psychology

Another sort of challenge to numerical degrees of confidence comes from Harman (1986) and Holton (2015). The basic idea of both authors (which others have supported) is that although it might be nice for some sort of epistemic agent to have degrees of confidence that obey something like the axioms of probability theory, it would be just too complicated for finite agents like us. Belief and intention are neither mere simplifications of an underlying confidence and expected value structure, nor are they supplements to it in special cases. Rather, they argue that most of our epistemology is constituted in full beliefs and intentions, with only a small role for degrees of confidence.

Bayesianism is not only a theory of the human mind. It is a theory of reasoning that is meant to be normative for humans, but also for groups like the scientific community as a whole, and for artificial or alien intelligences. This objection though, is targeted specifically at humans. It does not threaten Bayesianism as a theory of scientific reasoning for a community, or as a model for artificial intelligence, or as a conjecture about potential aliens we might some day encounter.

A full evaluation of the force of these arguments about humans would require detailed empirical investigation of the human mind that neither they nor I are in a position to carry out. My response to all of these arguments will take the form of arguing that the human mind has a lot more computational capacity than we consciously use, and this unconscious ability may well be sufficient to implement a sort of Bayesianism. Whether the human mind in fact does is an empirical question that will take a lot more research, but even if not I claim that a Bayesian mind would not be as different from the human mind as Harman and Holton suggest.

Harman's objection is that properly working with degrees of confidence is too difficult because of the number of different propositions one must work with. As I will describe in Chapter 6, the standard Bayesian picture for how one's degrees of confidence change over time requires that one have degrees of confidence in every conjunction of a proposition with potential evidence. If every proposition could possibly be learned, then that means that one needs to have a degree of confidence in every truth-functional combination of propositions one considers. Harman points out that this leads to a "combinatorial explosion" — with 10 basic propositions, there are 1024 truth-functional combinations, while with 20 there are over a million, and with 30 there are over a billion. He argues that having this many degrees of confidence would be far beyond the capacities of the human brain.

I think this claimed limitation on the human brain is quite unclear. The brain itself is quite complex, and well beyond our current understanding of it. We don't know what limitations it has. How many names do you know? Think of all the names of your friends and family, including "friends" that are merely connections on social media. How many more names do you know of celebrities, artists, philosophers, historical figures? Add to that all the people that you may have gone to school with, whose name you would recognize if you saw it in the right context, or local business owners that you'd recognize in combination with the name of their business. There are probably thousands, or even tens of thousands of such names. Is there any reason to believe that your brain is anywhere close to the limit of names you could know? With billions of neurons and who knows how many connections between them, there's no particular theoretical reason that it couldn't store simply astronomical amounts of information. Furthermore, nothing about the Bayesian idea requires these degrees of belief to be stored independently of each other — perhaps the brain works in some sort of compressed way, where degrees of belief in many related propositions are stored in some unified way that doesn't require each one to have its own separate representation. Stereotypes, biases, and heuristics are among the ways that we store many degrees of confidence with only a small amount of cognitive power.

While Harman thinks it a priori unlikely that humans *could* have minds powerful enough to deal with the large numbers of degrees of confidence involved in Bayesian reasoning, Holton thinks that there are reasons to believe that we *don't* have degrees of confidence (or "credences") of this probabilistic sort. He says, "the mental states involved here are nothing like credences. They do not typically register anything like numerical degree of probability; we cannot manipulate them; they do not obey anything like" the mathematical rules they are said to obey. He says, "It is possible that people had been entertaining and manipulating credences for millennia ... without realizing what they were doing, but it strikes me as implausible." The mathematics of probability theory is complex and difficult, and was only developed in the past few centuries. Holton suggests that therefore this can't be the sort of math that is going on in our minds.

I think this argument is based on a mistake like that of Zynda in arguing against strong realism. Both of these arguments seem to think that a strong realist view of an aspect of thought requires that we be aware of this aspect of thought. However, it's clear that there are many features of our mental life that we aren't directly aware of. Psychologists claim to have discovered many phenomena of the human mind — cognitive dissonance, confirmation bias, risk aversion, REM sleep. It seems quite plausible that people's thoughts had exemplified these phenomena for ages without realizing what they were doing.

This is true even for some quite complicated numerical phenomena. The human visual system is capable of quite complicated calculations. When trying to catch a ball thrown in a high arc, a person trying to catch the ball must predict where it will end up, which (according to Newtonian physics) involves computing a second-order differential equation, or at least tracing out a parabolic arc. However, people were able to catch thrown objects long before ancient mathematicians worked out the concept of a parabola, which itself happened long before early modern physicists argued that a parabola was the trajectory of a falling object. We don't need to be aware of a mathematical feature of our mind in order for it to work mathematically. This is perhaps even clearer for features like blood pressure, blood sugar, and neuron excitation, which can be observed to affect (or perhaps even constitute) thinking in a variety of ways, and are precisely numerical, but are very difficult[11] for people to introspect.

Holton further argues that not only are we not aware of any probabilistic reasoning, but in fact we do quite badly at it in many cases when it is made explicit. "If we are to be good subjects for the ascription of probabilistic attitudes, then we should be able to make the kinds of transitions — whether in beliefs or in behaviour — that would be expected. Yet a host of now very familiar research shows that, in many cases, we are very bad at this." (For many examples of this research, see Kahneman (2011).) The example Holton focuses on is the following:

> The probability that a woman has breast cancer is 1%. If she has breast cancer, the probability that a mammogram will show a positive result is 80%. If a woman does not have breast cancer, the probability of a positive result is 10%. What is the probability that a woman who has a positive mammogram result has breast cancer?

[11] Though not impossible — people often become aware of their "blood boiling" as blood pressure increases, or the sensation of getting "hangry" from low blood sugar.

As he notes, the mathematics of probability says that the probability of having breast cancer given a positive mammogram result in this case is less than 10%, but most people will answer that the probability is quite a bit higher.

However, this example seems to me irrelevant to the question of whether the mind works probabilistically. Consider a parallel case involving a physics problem:

> The velocity of a thrown baseball is 6 m/s at an angle of 60 degrees upwards from the horizontal. The acceleration due to gravity is 9.81 m/s^2 downwards. How far will the ball travel by the time it has fallen down to the same elevation from which it was thrown?

If you're anything like me, this question will be even harder to answer correctly than the question Holton considers. However, if we were actually to be in the situation described, rather than reasoning about it in the abstract, we would probably do pretty well at trying to catch the ball. It is extremely difficult to take a verbal description of probabilities, or velocities, and translate it into our credal or visual system, and then translate the answer we receive there back into numerical form. But this doesn't seem to me to give any reason to doubt the existence of some quantitative representation of these things in the mind.

When looking at a scene of people running around a playground, and throwing balls, one can see which people and objects are located closer or farther, which are taller or shorter, and which are moving faster or slower. But when asked to give numerical descriptions of the heights, distances, and speeds, one is unsure. Conversely, when looking at this scene and trying to imagine a 6 foot high person throwing a ball at 20 miles per hour, one is unsure quite how this would compare with everything else one sees. But if one in fact saw the person throwing that ball, it wouldn't be hard to catch. Similarly, when considering the uncertainties in your daily life, you have some sense of whether it's more likely that you will get stopped by a red light, or that the store will be out of yogurt, or that you will encounter your neighbor on the way out of the house. When asked to imagine a 20% chance of rain, you'll be unsure which of these events are more or less likely than it. But when presented with clouds with a particular distinctive look (that happens to be the look of 20% chance of rain in the next hour), one might do better.

The response I give here invites another challenge that is different from the one Holton makes. He raises worries about bad explicit reasoning with probabilities, which I think is no worry at all for there being a mental fact about the strength of confidence that is probabilistic. But there are also some well-known biases and errors that people make even when reasoning natively with confidence. People tend to feel a greater risk of crash when flying in a commercial airplane than they do when driving. And yet, as a matter of fact, there is a far greater risk of a crash when driving than flying.

This I think does point to a real inaccuracy in the confidence with which we hold various beliefs. But I think this is no more problematic than various visual illusions we are vulnerable to. Just because we are bad at turning our visual impressions of trees or buildings into correct verbal expressions of the height of those trees and buildings doesn't mean that our visual impressions fail to have quantitative structure that represents those heights in ways that behave like heights. Furthermore, in many cases, the visual and cognitive illusions we fall prey to are approximations that our mind uses to quickly and conveniently represent the world in ways that are usually fairly accurate, but have some systematic flaws. Our visual system has well-known illusions like the Müller-Lyer illusion, in which a line appears shorter when bookended with inward-pointing hooks, and longer when bookended with outward-pointing hooks. But as it turns out, in ordinary observations of three-dimensional scenes, lines bookended with inward-pointing markings *do* tend to be images of shorter objects than lines bookended with outward-pointing markings. (Howe and Purves, 2004) This is a useful heuristic that usually improves our estimate of lengths of three-dimensional objects, but tends to lead us astray when looking at two-dimensional drawings. Similarly, the various means by which we come to have greater or lesser degrees of confidence may do well in ordinary reasoning, even if they do badly in certain parts of the modern world with its novel media and information environments. (For more on these visual and cognitive illusions, see Chapter 8.)

A lot of mathematical sophistication is needed for doing detailed reasoning *about* degrees of confidence. But one doesn't need this mathematical sophistication to *have* them. One has a height, and a neural structure, and a whole cognitive structure of beliefs and desires with their logical relations to one another without having to be aware of it. Just as Molière's "Bourgeois Gentleman" was speaking prose his whole life without having realized it, I claim it is quite possible that we have been reasoning probabilistically with degrees of confidence our whole lives without realizing it. Our minds are in many ways quite a bit more sophisticated than we are explicitly aware of, and this may be another such way.

3 Imprecision

One last challenge to the claim that we really have degrees of confidence that are probabilistic is the idea that it is implausible that our degrees of confidence are quite so precise. There are a continuum of real numbers between 0 and 1, and it might seem odd that our finite minds can occupy those states so precisely. Many theorists have been motivated by the idea that although degrees of confidence might be *like* probabilities in some way, they can't be fine-grained enough to actually *be* probabilities, even in the attenuated sense suggested here.

Much recent philosophical literature on this idea suggests thinking of degrees of confidence not as individual real numbers, but instead as *sets* (usually intervals) of real numbers. (See, for instance Sturgeon (2008) and the reply by White (2009); Hájek and Smithson (2011); Elga (2010) and the reply by Moss (2015); and Schoenfield (2016).) Some of this literature recapitulates an earlier debate. (Levi, 1974, Jeffrey, 1984, Levi, 1985), in which Jeffrey suggests that it may be hard to pin down a precise number that represents a particular person's degrees of belief, while Levi suggests that degrees of belief may even behave in more “indeterminate” ways that don’t correspond to *any* number. There are several different motivations that may lead us to study theories of confidence that don’t work like precise real numbers. (Many such theories are described by Walley (1991).) But I will address one prominent style of arguing for this.

Miriam Schoenfield gives the following case:

> Detective Confuso
>
> You are a confused detective trying to figure out whether Smith or Jones committed the crime. You have an enormous body of evidence that you need to evaluate. Here is some of it: You know that 68 out of the 103 eyewitnesses claim that Smith did it but Jones’ footprints were found at the crime scene. Smith has an alibi, and Jones doesn’t. But Jones has a clear record while Smith has committed crimes in the past. The gun that killed the victim belonged to Smith. But the lie detector, which is accurate 71% percent of the time, suggests that Jones did it. After you have gotten all of this evidence, you have no idea who committed the crime. You are no more confident that Jones committed the crime than that Smith committed the crime, nor are you more confident that Smith committed the crime than that Jones committed the crime. (Schoenfield, 2012)

The important point she raises about this case is that it exhibits “insensitivity to mild sweetening”. If you were *equally* confident that Smith committed the crime as Jones, then if a single extra eyewitness testified that Smith did it, you would now lean towards Smith. However, from the construction of the case, it seems clear that you would be just as confused and exhibit the same indecision between the two even with the additional eyewitness.

However, I think that when we compare this sort of case to other sorts of features that we unhesitatingly report with precise real numbers, we can get a better understanding of how to think of them. Consider the height of a building. In 2009, the highest point in any building in the world was the tip of the Willis Tower (formerly Sears Tower). The Willis Tower is 527 m (1729 ft) high. However, if you look closely, the tip consists of two antennas that extend from the top, which had been replaced in 2000. If you only count the top of the architectural space, it is only 442 m (1450 ft), while Taipei 101 is 509 m (1671 ft) high. Looking more closely still, you’ll note that the top of Taipei 101 is an architectural spire that is not itself occupied. Its highest occupied floor is only at 438 m (1437 ft), while the Shanghai World Financial Center has its top floor at 474 m (1555 ft). (Willis Tower’s is only at 412 m (1354 ft).) Which building was actually the tallest in 2009? There’s a case to be made for each of them. Increasing or decreasing the height of any of these buildings by a few meters wouldn’t change this confusion and indecision.

This is one way I claim that indeterminacy about confidence could result. Just as it is vague which part of a physical object is part of a building for the purpose of determining its height, it could well be vague which aspect of her mental life is an agent’s degree of confidence in a proposition. (These were all surpassed by the Burj Khalifa in 2009, which is currently the tallest on all measures — tip and architectural roof at 828 m (2717 ft), and top floor at 584 m (1918 ft).)

Even once we have determined which features count for determining the height of a building, there is vagueness at a more fine-grained level. The steel, glass, concrete, paint, etc. that make up a building have surfaces that are constantly weathering and changing their structure at a molecular level. And the molecules themselves have indeterminate boundaries in their electron clouds (and at a much smaller scale, in their nuclei as well). The wavefunction may be defined in a perfectly determinate way, and features like heights can still be indeterminate.

However, on another level, the height of a tall building is variable over the course of a day, due to slight changes in size of metal spires as they heat and cool, and due to slight swaying in the wind (particularly when large storms come through). Thus, any report of the height of a building, if it is particularly precise, must be a report on the height at a particular moment in time, or the average height over a standard time period.

Similarly, degrees of confidence are presumably constantly fluctuating in real people. When considering the Detective Confuso case, I think it is natural to think that one goes back and forth as to which suspect is more likely to be guilty. This "going back and forth" might naturally be described by saying that one has degrees of confidence at each moment, but as one brings different aspects of the evidence to attention, one's degrees of confidence fluctuate up and down. Furthermore, shifts in blood sugar level, temperature of the room, blood pressure, and environmental factors, can often lead to shifts in attitude (regardless of whether or not these shifts are rational). Some of the fluctuation is due to attention to evidence, and other parts of it are due to factors that don't have any subjective presence. If this fluctuation is fast enough and large enough, it can give rise to cases that feel like the Detective Confuso case. At one phase in reasoning, one's degrees of confidence in each suspect being guilty fluctuate up and down past each other. After gaining some new evidence, it may be that one of them fluctuates around a slightly higher level than before, but they still pass each other quite regularly.

This fluctuation of confidence may be merely slightly variable, like the height of a building, or much more variable, like the depth of the water in an ocean with waves and tides, or the height of the trees in a forest as they sway in the breeze, or the height of the grasses in a field as they tremble in a storm. It may be slightly misleading to represent the height of a tree in centimeters, or the depth of the grasses in a field in millimeters. But there is in fact some fraction of a second in which one truly can say that the depth of the grass was exactly 543 mm (at least, up to the resolution of the surface of the solid molecules). Similarly, I claim that in many of these cases, it may well be true that at each moment, one has specific degrees of confidence that can be described with a precise real number, but since one's confidence is shifting so unsteadily, just a few seconds of thinking about the example can make it seem like there is no fact of the matter about which proposition one is more confident in.

Degrees of confidence may have some sort of vagueness of each of these sorts — they may be indeterminately constituted by components that themselves have indeterminacy (like whether the height of a building includes the spire), and they may have some sort of stable or unstable variance over short time periods (like the height of the grasses on a windy prairie). It may still be useful to describe them numerically in fairly precise ways, whether it is the way we say the Burj Khalifa is 828 m high, or only the way surfers might say that (as I write this), the wave forecast for Santa Cruz is 4-6 ft high today and 6-8 ft high two days from now. Setting up numerical degrees of belief by means of a representation theorem from measurement theory doesn't guarantee that they are more precise than the surf conditions at a time (which themselves are measured by meter sticks that satisfy the assumptions of their own measurement theoretic representation theorem), but it is surely open to empirical verification whether they are more like that, or more like the heights of buildings.

Research Questions

1. Is there a measurement theory for degrees of confidence and desire that can underlie the strong realist notions, or must one be a weak realist and construct them out of preferences?

2. Can one give a measurement theory for accuracy that will underlie the scoring rules used in Chapter 3?

3. Can a measurement-theoretic foundation for degrees of confirmation pick out the odds ratio, or are there different purposes served by different measures?

4. Can a weak realist about degree of confidence give a good measurement theory for degree of confirmation?

5. Does cutting-edge empirical psychology support or undercut the idea that the unconscious mind works probabilistically? (For an example of work that begins to address this question, see Goodman et al. (2015).)

6. Are there cases where the idea of precise numerical confidence breaks down in ways that can't be captured by the fluctuation picture?

References

Brennan, P., Crispo, A., Zaridze, D., Szeszenia-Dabrowska, N., Rudnai, P., Lissowska, J., Fabiánová, E., Mates, D., Bencko, V., Foretova, L., Janout, V., Fletcher, T., and Boffetta, P. (2006). High cumulative risk of lung cancer death among smokers and nonsmokers in central and eastern europe. *American Journal of Epidemiology*, 164(12):1233–1241.

Christensen, D. (1999). Measuring confirmation. *The Journal of Philosophy*, 96(9):437–461.

Churchland, P. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, 78:67–90.

Churchland, P. (1986). *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. MIT Press.

Clark, A. and Chalmers, D. (1998). The extended mind. *Analysis*, 58:10–23.

Cox, R. T. (1946). Probability, frequency and reasonable expectation. *American Journal of Physics*, 14(1):1–13.

Easwaran, K. (2015). Dr. Truthlove, or, how I learned to stop worrying and love Bayesian probability. *Noûs*.

Elga, A. (2010). Subjective probabilities should be sharp. *Philosophers' Imprint*, 10(5).

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. Technical report, The RAND Corporation.

Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science*, 66(3):S362–S378.

Gelman, A. and Nolan, D. (2002). You can load a die, but you can't bias a coin. *The American Statistician*, 56(4):308–311.

Good, I. J. (1985). Weight of evidence: a brief survey. In Bernardo, J., DeGroot, M., Lindley, D., and SMith, A., editors, *Bayesian Statistics 2*, pages 249–270. Elsevier.

Goodman, N., Tenenbaum, J., and Gerstenberg, T. (2015). Concepts in a probabilistic language of thought. In Margolis, E. and Laurence, S., editors, *The Conceptual Mind: New Directions in the Study of Concepts*. MIT Press.

Hájek, A. and Smithson, M. (2011). Rationality and indeterminate probabilities. *Synthese*, 187:33–48.

Halpern, J. (1999a). A counterexample to theorems of Cox and Fine. *Journal of Artificial Intelligence Research*, 10:76–85.

Halpern, J. (1999b). Cox's theorem revisited. *Journal of Artificial Intelligence Research*, 11:429–435.

Harman, G. (1986). *Change in View: Principles of Reasoning*. MIT Press.

Holton, R. (2015). Intention as a model for belief. In Vargas, M. and Yaffe, G., editors, *Rational and social agency: Essays on the philosophy of Michael Bratman*. Oxford University Press.

Hosiasson-Lindenbaum, J. (1940). On confirmation. *Journal of Symbolic Logic*, 5(4):133–148.

Howe, C. and Purves, D. (2004). The Müller-Lyer illusion explained by the statistics of image-source relationships. *Proceedings of the National Academy of Sciences*, 102(4):1234–1239.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.

Jeffrey, R. (1984). Bayesianism with a human face. In Earman, J., editor, *Testing Scientific Theories*, volume 10 of *Minnesota Studies in the Philosophy of Science*, pages 133–156. University of Minnesota Press.

Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

Krantz, D., Luce, D., Suppes, P., and Tversky, A. (1971). *Foundations of Measurement*, volume 1. New York Academic Press.

Levi, I. (1974). On indeterminate probabilities. *The Journal of Philosophy*, 71:391–418.

Levi, I. (1985). Imprecision and indeterminacy in probability judgment. *Philosophy of Science*, 52:390–409.

Lindley, D. (1982). Scoring rules and the inevitability of probability. *International Statistical Review*, 50(1):1–11.

Lloyd-Jones, D., Leip, E., Larson, M., D'Agostino, R., Beiser, A., Wilson, P., Wolf, P., and Levy, D. (2006). Prediction of lifetime risk for cardiovascular disease by risk factor burden at 50 years of age. *Circulation*, 113(6):791–798.

Meacham, C. and Weisberg, J. (2011). Representation theorems and the foundations of decision theory. *Australasian Journal of Philosophy*, 89(4):641–663.

Moss, S. (2015). Credal dilemmas. *Noûs*, 49(4):665–683.

Nash, L. (1974). *Elements of Statistical Thermodynamics*. Addison-Wesley.

Putnam, H. (1975). The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science*, 7:131–193.

Ramsey, W. (2013). Eliminative materialism. *Stanford Encyclopedia of Philosophy*.

Schoenfield, M. (2012). Chilling out on epistemic rationality: A defense of imprecise credences (and other imprecise doxastic attitudes). *Philosophical Studies*, 158(2):197–219.

Schoenfield, M. (2016). The accuracy and rationality of imprecise credences. *Noûs*.

Sellars, W. (1956). Empiricism and the philosophy of mind. In Feigl, H. and Scriven, M., editors, *Minnesota Studies in the Philosophy of Science, Volume I: The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, pages 253–329. University of Minnesota Press.

Sturgeon, S. (2008). Reason and the grain of belief. *Noûs*, 42:139–165.

van Fraassen, B. (1980). *The Scientific Image*. Clarendon Press.

Vassend, O. (2015). Confirmation measures and sensitivity. *Philosophy of Science*, 82(5):892–904.

Vassend, O. (2018). Goals and the informativeness of prior probabilities. *Erkenntnis*, 83(4):647–670.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall.

White, R. (2009). Evidential symmetry and mushy credence. *Oxford Studies in Epistemology*.

Zynda, L. (2000). Representation theorems and realism about degrees of belief. *Philosophy of Science*, 67:45–69.