An Opinionated Introduction to the Foundations of Bayesian Epistemology

Kenny Easwaran

September 7, 2015

1 Belief comes in degrees

Epistemology is the study of knowledge and related states. Historically, much western epistemology has focused on the notion of belief and what it takes for a belief to be justified. [CITATIONS] More recently, there has been greater discussion of interpersonal issues in epistemology (such as trust, testimony, disagreement). [CITATIONS] Formal epistemology is the application of mathematical methods to these issues. There are a great number of different projects that have been pursued by these means, including the use of “epistemic logic” to understand knowledge [CITATIONS], “doxastic logic” to understand belief [CITATIONS], and many different projects in formal social epistemology [CITATIONS].

The focus of this book is Bayesian epistemology, which uses the mathematics of probability theory to help us understand the phenomenon of confidence, which is taken to be a state like belief that comes in greater or lesser degrees. (The term “Bayesian” is derived from a paper published by the Rev. Thomas Bayes. For more on whether this terminology is accurate, see Earman (1992). [Say something about the origin of the term in statistics.]) The goal of this book is to develop a picture of this phenomenon of confidence more clearly, in order to see the ways in which it is appropriate to use the mathematics of probability to help explicate it, and to explain the normative requirements that apply to it.

Many of the authors discussed in these first three chapters assume that confidence can be represented numerically. However, I will focus in these early chapters on the arguments that seem to apply more generally. Numerical representation will only enter in Chapter 4. The important thing in these introductory chapters is just that confidence is a phenomenon that can come in greater and lesser degrees.

1.1 What is belief?

Belief, as philosophers mean it, is a state of affirmation or judgment or the like. Although ordinary English speakers often reserve the verb “believe” for beliefs of a spiritual or ethical nature, philosophers intend to include the states ordinarily
ascribed with the verb “think” or “know”. However, philosophers aren’t the only ones to use the word in this way. Note the famous line from W.C. Fields playing on these two uses — “Everybody’s got to believe in something. I believe I’ll have another beer.” Confidence is in many ways related to belief, so to figure out what confidence is like, it will be useful to consider what belief is like.

On some level this is an empirical question for psychology. Belief is a mental state belonging to humans, and so the scientific study of human minds is the best way to find out what this state is like. And much work in recent decades has upended older thoughts about belief. A certain sort of Cartesian psychology might have assumed that the contents of one’s own mind are transparent to introspection, so that one is always aware of one’s beliefs. But it’s clear that much about the mind is opaque to introspection. Illusions, implicit bias, and various other effects make it clear that we aren’t always aware of what our experiences are like or what our overall attitudes are like. Tamar Gendler has discussed cases where our behaviors contradict our avowed beliefs, as when a person who consciously knows she is safe behind glass finds herself unable to walk to the edge of the glass observation deck on a tall skyscraper. (Gendler calls these states “a-liefs” in contrast with “be-liefs”.) Discussion of this sort of opacity of the mind goes back at least as far as Sigmund Freud’s discussion of the workings of the unconscious, though modern scientific methods are able to reveal much more about what they are like.

While there is surely much to be gained from this sort of empirical study of the mind, this is not the sort of question that I am most interested in. Rather, I am interested in a more general (and perhaps abstract) characterization of a certain type of idealized reasoner. (More about the particular type of idealization involved will come in Chapter 6.)

Rather than asking what sort of mental states we actually have many philosophers have asked questions about the mental states of idealized reasoners. A sort of Kantian question would be, what sorts of mental states are required in order for someone to count as a reasoner at all? This may then tell us about the nature of rationality, if maintaining mental states of that sort is required in order to continue to count as a reasoner. A sort of Carnapian question would be, what sorts of mental states would be useful for someone to have in order to achieve various goals? This would tell us about rationality in a different sort of sense. (This Carnapian project could further be pursued by someone uninterested in rationality in general, as part of, for instance, a project to develop versions of artificial intelligence.)

Both of these projects need some understanding of the purpose of having beliefs. As a psychological question, this could be answered by understanding the biological function of mental states of various sorts. To address this biological question, we would investigate what roles that mental states play are important enough to shape the way they persist in the species over time. But again, I am interested in the question of what purpose beliefs need to serve in order to count as beliefs, that is, to play a certain role in the life of a reasoner.

There can be instrumental goals or intrinsic goals to the idea of belief. The most common instrumental goal discussed has been the goal of guiding actions
in a useful way. Hume’s theory of motivation suggests that intentional actions are guided by the beliefs and desires of an agent, acting together. Roughly, an agent will do action $A$ if she believes that $A$ will best satisfy her desires. Beliefs about one part of the world, together with desires about a related part, will result in action. Understanding this connection will help us understand what beliefs should be like. This idea is explored in Chapter 2.

The most common intrinsic goal that has been discussed is the goal of correctly describing the world. The idea is that belief is a mental state that could exist even in isolation from desire and action, but that it has an aim of correctness. Understanding the implications of this view is the goal of Chapter 3. One interesting thing to note though is that both belief and desire aim to correctly describe the world, though they do so in different ways. As Anscombe noted in her book *Intention* [CITATION], although belief and desire both aim to correctly describe the world, they do so with opposite “directions of fit”. When the belief and the world are mismatched, one aims to change the belief. When the desire and the world are mismatched, one aims to change the world. Desire and belief are both in some sense representational states (in that they represent the world as being a certain way), but belief has a mind-to-world direction of fit while desire has a world-to-mind direction of fit.

I won’t say much about what it takes for a mental state to be “representational”. (I take it that some sort of teleosemantic picture can be made to work, along the lines of Ruth Millikan or Karen Neander, but this point is beyond the scope of my project.)

1.2 What are the objects of belief?

This characterization of belief as a representational mental state with a mind-to-world direction of fit leaves further questions open about the ways in which belief represents the world. The totality of this representational mind-to-world attitude is what I will call an agent’s “doxastic state”. The doxastic state is surely a very complex whole in actual humans, and may not be separable in any way. However, in ordinary language, we tend to report on each other’s doxastic state by in a sense “factoring” it into attitudes towards various particular claims, rather than to the world as a whole. For instance, we might say, “Lois thinks Clark is at work” to indicate that Lois’ mental state represents the world as being such that Clark is at work. Further to this, we might suspect that the entire doxastic state could hypothetically be described in terms of attitudes of this sort to individual contents.

We might ask several questions about this sort of factorization. First, is there some deep sense in which this is a good account of the doxastic state, or is it merely a theoretical convenience? Second, what is the nature of these contents that make this account of the doxastic state plausible or useful? Third, how many different attitudes are involved in this factorization, and how are they related to each other? The first two questions are the target of this section, and the third question is the target of the next section.

The question of how actual human doxastic states factor is of course an
empirical question. Nevertheless, some philosophers have made proposals of this sort. One such proposal is the view associated with Jerry Fodor, saying that the human mind is composed of many attitude “modules” that implement its many different functions. These modules are in constant communication with each other, and have some sort of representation akin to a language in which signals are passed back and forth. Fodor and others often put the view very crudely, and call one of these modules the “belief box”, and say that believing a claim just is having a “sentence written in mentalese” expressing that claim that is present in the belief box. (In some sense, this is just a modern rendition of Frege’s view that mental attitudes involve grasping abstract “senses” and “thoughts” — Fodor has just proposed that the abstract compositional structure of Fregean thoughts is represented by similar structure of mental contents.)

Some others have followed Wittgenstein in rejecting the possibility of this sort of “private language”, but have instead taken sentences in a public language to be the relevant contents. Following some sort of behaviorist or operationalist practice, one might then attempt to characterize the doxastic state in terms of responses to these sentences. The relevant responses might just be dispositions to bet, or might involve more complex and sophisticated reasoning and decision-making behavior. Some version of this view seems to have been characteristic of Donald Davidson, and perhaps David Lewis. Whether this factorization is taken to be a real description of the psychological states involved, or just a theoretical convenience for using them, may depend on whether these dispositions are sufficiently stable. Some research in social psychology seems to indicate that seemingly irrelevant “priming” can drastically change one’s responses, so that there may not be such real dispositions, but this literature is notoriously fraught with irreproducible experiments. But we could also consider that whether or not people actually have such dispositions, it would be useful for us to, so that some sort of idealized rational agents would be characterized well in this way, even if actual humans are not. (We will return to questions about the relevant sort of idealization in Chapter 7(?).)

Another way of thinking about doxastic states has taken as its basis the idea that fundamentally, representations of the world involve considerations of various ways that the world might be. These “ways the world might be” can then be thought of in various ways. [CITATIONS - Leibniz, Barcan Marcus, Kripke, Nolan, Jago?]

One particular version of this view is associated with much of the work of Bob Stalnaker. A doxastic state is thought of as a set of possible ways the world might be. To say that someone believes something is just to say that her set of doxastic possibilities are all ones on which that content is true. We can fully describe a doxastic state of this sort by listing propositions that the agent does or doesn’t believe, but importantly we don’t need to list them all. On this view, beliefs are automatically closed under logical consequence, and are automatically consistent. Of course, this sounds implausible for ordinary descriptions of people’s mental states — we are used to saying that people might be inconsistent, and may fail to recognize consequences of some of their thoughts. But on the Stalnaker view, that is all a surface phenomenon masking
Although doxastic states of this sort can be described in a factorized way, this is merely a potential convenience. In many ways it is better to describe a doxastic state just as the set of possibilities. Whether or not this is an accurate characterization of the Stalnaker view, or whether it is a good account of the mind, it helps illustrate ways that a factorization of a doxastic state might relate to the state as a whole.

This crude picture gives very implausible results if we think of the possibilities involved as being mere possible worlds. For instance, necessary truths are automatically believed by everyone. This includes both the fact that $2 + 2 = 4$ and the fact that $167 \cdot 263 = 70,221$. But some of the implausibility of this view can be mitigated if we take the possibilities to have additional features. We might follow Stalnaker’s version of “two-dimensional semantics” — roughly, although along one axis of variation, the world can’t change in a way to make “$167 \cdot 263 = 70,221$” come out false, the use of language can, and a possible world in which language is used slightly differently can represent the claim as being false, so that uncertainty about whether that possibility is actual represents the agent’s uncertainty about whether $167 \cdot 263$ really is $70,221$. Or we might say instead that the sort of possibilities involved here are a special sort of “doxastic possibilities” that aren’t limited by what is metaphysically or even logically possible.

I generally think a version of this sort of view will be most useful. There may in fact be sentences (in public language or mentalese), or abstract structured propositions. But the only feature they have that I suspect is relevant to one’s doxastic attitudes towards them is the set of doxastic possibilities in which they are true. Thus, much of this book will be written as though the objects of mental attitudes just are sets of doxastic possibilities. They may in fact have further structure to them, but I mean to suggest that this set structure is all that is essential for most of the uses we put them to.

### 1.3 How is confidence related to belief?

If a doxastic state can be described in terms of sets of possibilities of some sort, the question remains what attitudes are part of the doxastic state. The Stalnaker picture mentioned above suggests that possibilities are either in or out, so the natural way to think of the attitude is as one that is either present or absent. From now on, I will use the word “belief” to refer to that sort of attitude. And yet there also seems to be a phenomenon of being more or less confident in some possibilities. I have decreasing confidence in the following six propositions: that the sun will rise tomorrow; that I will make it to class this afternoon; that it will rain later this week; that there is life on Mars; that humans evolved from river-dwelling apes; that I don’t have two hands. I believe both of the first two, disbelieve the last two, and neither believe nor disbelieve the middle two. Thus, it appears that both belief and confidence are useful ways to describe some of my attitudes. (I will avoid considering the question of whether there is some third attitude that also constitutes part of a doxastic state. I
don’t think that “supposing” or “entertaining” or “considering” a proposition count as doxastic attitudes.)

Given this appearance, there are several possibilities. It could be that one of these apparent states is just an illusion, caused by our own lack of introspective abilities. Or on the more philosophically idealized approach, perhaps only one of these attitudes is essential to being a good reasoner, and the other is just a feature that humans contingently have as well. Another subtly different possibility is that although both of these attitudes are correct descriptions of part of our psychology (or the psychology of an idealized agent), but that one of them is just a summary of certain facts about the other, so that only one is fundamental. (Compare: evolution proceeds both by Darwinian natural selection and Mendelian genetics, but the Mendelian genes are the constituents of Darwinian inherited characteristics — in the early 20th century these were seen as competing theories until one was reduced to the other.) Another subtle difference: it could be that both are accurate descriptions in their own way of a single underlying doxastic state — neither is more accurate than the other, but they are not independent either. Finally, it could turn out that the initial description is correct, and belief and confidence are just two separate aspects of our doxastic state.

1.3.1 Belief reduces to confidence

Among philosophers who have explicitly considered this question, the most popular view has probably been that confidence is the more fundamental notion. Bruno de Finetti (1931) suggests that the goal of a doxastic state is just to make predictions about the world. Due to the skeptical considerations of Descartes and Hume, he thinks that these predictions must always come with varying degrees of uncertainty. Thus, even if we do have an attitude of belief, it is unnecessary, and more idealized reasoners would just work with greater or lesser degrees of confidence.

Frank Ramsey (1926) comes to a similar view for behaviorist reasons. He thinks that the way to understand doxastic states is through behavior, and he thinks that the only doxastic states that can be understood through behavior are states of greater or lesser confidence. Thus, confidence is the only real doxastic state. Richard Jeffrey (1970) says that Ramsey’s analysis of confidence “sucked the marrow out of the ordinary notion, and used it to nourish a more adequate view.” (p. 172) Even if belief does exist beyond confidence, it is confidence that explains all the interesting doxastic phenomena.

Some more recent philosophers have taken less radical versions of this view. Richard Foley (1993) claims that confidence and belief both exist, but that belief just is sufficiently high confidence. (By a simplistic reading of a certain passage from John Locke, this view has come to be known as the “Lockean thesis”.) That is, belief is not a significantly different state than confidence, but is just a rough “rounding off” of confidence that we often talk about because it is simpler to discuss. An important and interesting consequence of this view is that belief is not necessarily closed under conjunction, which allows Foley to make simple
sense of the Preface Paradox. (Makinson, 1965) One can be quite confident of each claim one has made in a large book, and yet also be quite confident of the claim that one makes in the preface — that there is still an error somewhere in the book. However, the parallel solution seems less plausible in the formally similar lottery paradox. (Kyburg, 1961) No matter how many tickets there are in a lottery, it doesn’t seem correct to say that one thinks, of each ticket, that it will lose. More sophisticated reductions of belief to confidence have been developed that try to avoid these issues. (Arló Costa and Pedersen, 2012, Lin and Kelly, 2012, Leitgeb, 2014)

1.3.2 Confidence reduces to belief

The converse view has also had various defenders over the years. A very forceful argument of this sort has been given by Gilbert Harman (1986). Harman actually argues that confidence itself is too complex to be a central part of our doxastic state. Once one has formed some beliefs, one can in a sense discard all the possibilities in which they are false. Having a greater number of beliefs can in some sense simplify one’s cognitive processing. However, with confidence, one never rules anything out, and one has to keep track of greater and lesser degrees of confidence for very many propositions. Harman makes a very specific argument that if one adds just a few more propositions to consideration, then one gets an exponential increase in the number of possibilities that one must track. Adding a new proposition doesn’t just add two new objects of confidence — it doubles the total number of possibilities, because all existing possibilities must be considered both in a form where the new proposition is true and in one where it is false.

Of course, the success of this argument depends on empirical facts about the processing power of the human brain. It could well be that humans are actually capable of far more complex mathematical processing than we are aware of. It is often easy for a human to catch a ball that is thrown, even though computers have great difficulty solving the three-dimensional differential equations needed to track a moving object. And there are a great many other mental modules that seem to do computationally very difficult tasks — we can recognize faces, maintain our balance while walking, and fold a towel, all without apparent effort. These capabilities have only very recently been demonstrated in computers, operating at the limits of current computational ability. Perhaps the human brain really is powerful enough to keep track of the things that Harman thinks are impossible.

At any rate, there are also philosophers who have defended the view that belief is the more fundamental notion, so that our confidences always reduce to belief, rather than being independent states that must be tracked. The most popular way to do this has generally involved the idea that confidence is really belief in a separate sort of content. Consider the fact that I am more confident that it will rain later this week than that there is life on Mars. It is common to express this confidence by saying something like, “I think it’s more likely that it will rain this week than that there is life on Mars.” Many philosophers
have used the phrase “think it’s more likely” to suggest that this is really a belief about a sort of objective probability, rather than a subjective degree of confidence.

Making sense of the relevant notion of objective probability has been difficult however. Although I might believe that the weather is fundamentally indeterministic, so that there is some objective chance that rain will occur and some chance that it won’t, it seems implausible that I think the status of life on Mars is the same. Surely there is already a fact there, and I don’t think there is any objective indeterminacy. Thus, if there is a notion of objective probability here, it must be something more like “evidential probability”. I must think that somehow my evidence objectively supports rain this week more than it supports life on Mars. Of course, whether or not one can defend such a notion of objective evidential support, it seems strange to think that having greater or lesser confidence requires having thoughts about the objective strength of evidence.

Another way to defend the fundamentality of belief over confidence would say that “think it’s more likely” doesn’t require a notion of “likeliness” other than confidence, and then try to explain confidence in some other way in terms of belief. In my (Easwaran (ms)) I put forward a view of this sort, though I am not very confident that it is correct. The idea (which will be discussed in a bit more detail in Chapter 3) is that confidence is not a property of an individual belief, but somehow arises from the overall pattern of one’s beliefs. To be more confident in $p$ than in $q$ just is to believe more propositions that are logically connected to $p$ in positive ways than to $q$. In that paper I argue that there is a way to make sense of this view that actually gives rise to the standard mathematics of probability theory that has often been used to make sense of confidence. But for our purposes at this point, it suffices if this can give rise to a comparative notion at all.

1.3.3 Belief and confidence both exist

Finally, one could claim that both belief and confidence are real doxastic phenomena, and neither can be explained entirely in terms of the other. (For the purposes of this book, I am ignoring the possibility that both descriptions of a doxastic state in terms of attitudes are on the wrong track, though this is surely a serious possibility to consider as well.) This view has become more popular in recent years, now that belief and confidence are both well-established notions in epistemology.

One family of views of this sort has been defended in various forms by Weatherson (2005), Ross and Schroeder (2011) and others. Just as Ramsey and others suggested that confidence is a disposition to act in ways that run certain risks but not others (to be discussed more in Chapter 2), this view says that to believe that $p$ is to “treat $p$ as true in practical reasoning”. That is, when one believes that $p$, one acts in ways that would make sense only if $p$ were true. Weatherson suggests that there might be a way to reduce this to confidence, so that belief just is confidence above whatever threshold is relevant in the practical scenarios one actually finds oneself in. This allows that a change
in one’s practical scenario can result in a change in belief, with no change in confidence or evidence. For instance, one might currently believe that the bank will be open on Saturday, but if one suddenly learns that an important financial transaction must be made before Monday, one’s confidence in the bank being open Saturday may not be sufficient to warrant putting off the transaction until then. The presence of the pressing financial circumstance changes the threshold needed for confidence to constitute belief.

Ross and Schroeder, on the other hand, argue that the confidence by itself doesn’t determine whether one’s attitude constitutes belief. Even in the relatively low stakes scenario, someone who reasons by explicitly considering the possibility that the bank might be closed exhibits different dispositions from someone who just ignores the possibility, even if they have equal confidences. Although the confidence might be high enough to justify believing that the bank will be open, whether the person actually has this belief is a separate question.

A very different way of thinking of belief has been proposed by Justin Dallmann. (2015) Dallmann’s proposal takes Harman’s objection seriously. Correct reasoning with confidence surely is more complex than correct reasoning with beliefs. However, Dallmann suggests that there are certain cognitive shortcuts that might be relevant for reasoning with confidence. Although keeping track of the precise degree of confidence one is warranted in having by one’s evidence is extremely difficult, for many propositions it is clear that the relevance of new evidence that one is gathering is quickly diminishing. After gathering enough evidence to support a proposition to a very great degree, it may be better to just commit to keeping that confidence high, and focusing one’s cognitive efforts on evaluating the relevance of new evidence to other propositions that one may be interested in. Dallmann suggests that this commitment to stability just is what we call belief. Belief is a sort of shortcut around reasoning with confidence, and computationally limited agents like us must use it in order to keep our cognitive load manageable.

Both of these sorts of views that allow for both confidence and belief raise important questions about the relation between the two. If neither is constituted by the other, then it seems that it should be possible for an agent to get into some very strange states — perhaps she can believe something while being extremely confident that it is false. If this is impossible, then these theorists need to say why it is impossible in a way that doesn’t make one of the two states entirely reduce to the other. More likely, these theorists will just say that it is irrational to believe something while being extremely confident that it is false, and explain the lack of such states by suggesting that humans are generally relatively rational.

Although the form of such rational requirements will be of greater centrality in Chapter 7, there are a few interesting features of such requirements that are worth considering at this point. John Broome has drawn a distinction between “narrow scope” and “wide scope” requirements. Narrow scope requirements here could be framed in two ways — we could say that if one believes that \( p \), then one would be irrational in having very low confidence that \( p \), or conversely we could say that if one has very low confidence that \( p \), then one would be
irrational in believing that \( p \). Although these two requirements both describe the overall state of believing that \( p \) while having low confidence in it as irrational, they put the irrationality on quite different features of one’s doxastic state.

However, either version of this rule seems to raise the worry that one can put rational requirements on oneself by having a doxastic state with certain features. Although one sort of attitude may be answerable to the world, the other seems to be answerable to the first rather than to the world directly. Broome and others have suggested that it is more plausible for these norms of rationality to take “wide scope” — it is rationally required to be such that one doesn’t both believe that \( p \) while having low confidence in \( p \). This way of putting it doesn’t say which of the two attitudes is the source of irrationality, and even allows that both might be.

Niko Kolodny has argued, however, that such wide-scope norms are never fundamental. They must always follow from narrow scope ones. Perhaps a plausible way for this to come about would be to endorse a sort of “evidentialism”. One might think that there is a requirement saying that, “if you don’t have strong evidence in favor of \( p \), then it would be irrational to believe that \( p \)”, and another requirement saying that, “if you do have strong evidence in favor of \( p \), then it would be irrational to have very low confidence in \( p \)”. With these two norms together, we can then see that, regardless of one’s evidence for \( p \), either it is strong (in which case it is irrational to have low confidence in \( p \)) or not (in which case it is irrational to believe that \( p \)), and in no circumstance can it be rational to both believe that \( p \) and have low confidence in \( p \). If something like this view is right, then the wide-scope norm relating belief and confidence is merely a consequence of more fundamental narrow-scope norms relating evidence separately to both belief and confidence.

While I don’t happen to believe that this sort of explanation is likely to work out (I have many worries about the concept of “evidence” and the ways that it can provide norms on belief and confidence), I do think that something of this form is more promising than narrow-scope norms directly relating belief and confidence.

1.4 Conclusion

The object of study of this book is the doxastic state of some sort of idealized reasoning agent. I claim that there is some useful way to think of a doxastic state in terms of one or more attitudes that one may have to various contents. Among the attitudes that is of interest (whether fundamental or not) is an attitude of confidence that comes in degrees. The objects of this attitude can usefully be thought of in terms of sets of doxastic possibilities. There may further be a notion of belief, which may be more or less fundamental than confidence, or just a separate aspect of our doxastic state. The objects of belief and confidence may have some sort of linguistic or conceptual structure to them. But my central claim is that there is a notion of confidence, and the fundamental notions of rationality apply to confidence in terms of the possibilities in which its objects are true or false.
References


Easwaran, K. (ms). Dr. Truthlove, or, how i learned to stop worrying and love Bayesian probability. unpublished.


