Updating, Undermining, and Perceptual Learning

Brian Miller

July 3, 2015

Abstract

As I head home from work, I’m not sure whether my daughter’s new bike is green, and I’m also not sure whether I’m on drugs that distort my color perception. One thing that I am sure about is that my attitudes towards those possibilities are evidentially independent of one another, in the sense that changing my confidence in one shouldn’t affect my confidence in the other. When I get home and see the bike it looks green, so I increase my confidence that it is green. But something else has changed: now an increase in my confidence that I’m on color-drugs would undermine my confidence that the bike is green. Jonathan Weisberg and Jim Pryor argue that the preceding story is problematic for standard Bayesian accounts of perceptual learning. Due to the ‘rigidity’ of Jeffrey conditionalization, a negative probabilistic correlation between two propositions cannot be introduced by updating on one of them. Hence if my beliefs about my own color-sobriety start out independent of my beliefs about the color of the bike then they must remain independent after I have my perceptual experience and update accordingly. Weisberg takes this to be a reason to reject Jeffrey conditionalization. I argue that this conclusion is too pessimistic: conditionalization is only part of the Bayesian story of perceptual learning, and the other part needn’t preserve independence. Hence Bayesian accounts of perceptual learning are perfectly consistent with potential undercutters for perceptual beliefs.
1 Introduction

On the way home from work I find myself wondering what color my daughter’s new bike is. I think it might be blue, or red, or maybe green — I’m not sure. I’m also not sure whether my coworker was joking when he claimed to have slipped a (slow acting) color-hallucination-inducing drug in my afternoon coffee. One thing I am sure about at this point is that facts about my perceptual sobriety and facts about the color of my daughter’s bike are evidentially unrelated: since I haven’t yet seen the bike, changing my confidence in the one shouldn’t affect my confidence in the other. Later I see the bike, and since appears to be green I become confident that it is green. But something else has changed as well: now if I were to increase my confidence that I’m on color-drugs, I would begin to doubt the veridicality of my perceptual experience as of the greenness of the bike, and as a result I would reduce my confidence that the bike is green. In other words, my belief about whether I’m on color-drugs is no longer evidentially unrelated to my belief about the color of the bike; the former now serves as a potential defeater for the latter. In particular, it’s an undermining defeater: instead of telling directly against the truth of the bike is green, it tells against the evidential support that I have for believing that proposition.

Jonathan Weisberg (2009), (2014) and Jim Pryor (2013) have argued that the case as described is in tension with the Jeffrey Bayesian’s account of perceptual learning. That’s because any two propositions that start out probabilistically independent cannot lose that independence as a result of conditionalizing on one of them. Conditionalization being the primary means of rationally permissible credence revision in any Bayesian account of perceptual learning, and the loss of probabilistic independence being essential to at least some cases of undermining defeat, they conclude that undermining defeat and Jeffrey Bayesianism are in tension or even inconsistent.\footnote{Thought both Pryor’s and Weisberg’s written work supports this reading, in conversation they both take the lesson of the puzzle to be somewhat weaker: Weisberg takes it as a reason to abandon subjective Bayesianism for an objective version that permits conditionalization directly upon perceptual states, while Pryor takes the lesson to be very} In this essay I argue that Weisberg’s and Pryor’s conclusion is
overly pessimistic, and that the Bayesian account of perceptual learning is perfectly consistent with undermining defeat.

2 The Puzzle

I’ll begin with a quick sketch of the Bayesian account of perceptual learning that I’ll be discussing: agents assign subjective probabilities or credences to propositions (e.g. \( P(A) \)), with those assignments subject to norms of probabilistic coherence (call that thesis ‘Probabilism’). Probabilities are also assigned to propositions conditional on other propositions (e.g. \( P(A|B) \)), which for our purposes I’ll understand as being defined in terms of unconditional probabilities according to the formula

\[
P(A|B) = \frac{P(A \& B)}{P(B)}.
\]

Perceptual experience leads agents to revise some subset of their probability assignments, which by a process of Bayesian conditionalization leads to revisions in other probabilities.

Bayesians understand the process of conditionalization in slightly different ways. According to Classical Bayesians, upon changing the probability in \( B \) to 1 (due to a perceptual experience, or whatever) the agent updates by setting her new credence in \( A \) to her old credence in: \( A \) conditi \( B \). In other words, where \( P_{\text{old}}(.) \) is the probability function accepted by the agent before having the relevant perceptual experience and \( P_{\text{new}}(.) \) is the function accepted after the experience and updating on \( B \), Classical Bayesians claim that for any \( A \) and \( B \),

**Classical conditionalization**: \( P_{\text{new}}(A) = P_{\text{old}}(A|B) \)

Jeffrey Bayesians\(^2\) generalize the Classical program by relaxing the requirement that all conditionalization be on propositions assigned a credence of 1. I’ll go into more detail about how Jeffrey conditionalization works similar to what I argue below. In this essay I’ll be responding to their written work.

\(^2\)For the purposes of this paper a Jeffrey Bayesian is any Bayesian who accepts Richard Jeffrey’s generalization of the classical rule of conditionalization that I’m calling ‘Jeffrey conditionalization’. Our ‘Jeffrey Bayesians’ needn’t share Richard Jeffrey’s particular views about the motivations for accepting that rule (Jeffrey (1992)), Radical Probabilism (Jeffrey (2004)) or anything else.
below, but here’s a rough sketch: the process begins with an assignment of credences to some subset of the propositions to which the agent assigns credences. These credence assignments are laundered (see below) into a partition of the agent’s state space, in which that space is divided into an exhaustive and exclusive set of ways that the world might be (the ‘elements’ of the partition), with each way assigned a credence. Finally, the agent conditionalizes on this partition with its weighted elements — call them $B_i$ — using the following rule:

**Jeffrey conditionalization:**

$$P_{\text{new}}(A) = \sum_i P_{\text{old}}(A|B_i)P_{\text{new}}(B_i)$$

With these preliminaries in place, let’s return to the puzzle of the color-drugs and the bike. Before seeing the bike I regarded the veridicality of my own color perception as irrelevant to the greenness of the bike, and hence I regarded the propositions *I’m on color-drugs* and *the bike is green* as being probabilistically independent. Taking $P_{\text{old}}$ as the probability function that I accepted before having a perceptual experience as of the greenness of the bike, that means that:

(1) $P_{\text{old}}(\text{the bike is green }| \text{I’m on color-drugs}) = P_{\text{old}}(\text{the bike is green})$

After I’ve had an experience as of the bike being green and shifted my partition accordingly, I adopt the credence function $P_{\text{new}}$ that results from the relevant conditionalization procedure. At this point I no longer regard the two propositions as being independent, but instead regard *I’m on color-drugs* as a defeater for *the bike is green*. I’ll interpret this as saying that:

(2) $P_{\text{new}}(\text{the bike is green }| \text{I’m on color-drugs}) < P_{\text{new}}(\text{the bike is green})$

Weisberg and Pryor observe that the introduction of a negative probabilistic correlation between two propositions through a process of updating on one of them is problematic within the Bayesian framework. That’s because Jeffrey conditionalization is rigid:\footnote{See Jeffrey (1992, p. 80) and Weisberg (2014, p. 125).}
**Jeffrey Rigidity:** If $P_{new}$ is the credence function resulting from accepting $P_{old}$ and then updating on a shift in partition \( \{B_i\} \), then for any proposition $A$ and any $B_i \in \{B_i\}$, $P_{new}(A|B_i) = P_{old}(A|B_i)$

and rigid updating rules preserve independence:4

**Rigidity is Independence Preserving (RIP):** If the transition from $P_{old}$ to $P_{new}$ is rigid on partition $\{B_i\}$ and $P_{old}(B_i|A) = P_{old}(B_i)$ for every $B_i$, then for every $B_i$, $P_{new}(B_i|A) = P_{new}(B_i)$ for every $B_i$

Hence Weisberg’s puzzle, as I’ll call it, is this: our intuitions about undermining defeaters commit us to both (1) and (2), but due to the rigidity of Bayesian conditionalization, when we update on *the bike is green* that combination is impossible.5 As Weisberg puts it, “...[perceptual undercutters are] irrelevant to the supported proposition at first, but negatively relevant after the perceptual state has lent its support. And this is precisely what ‘Rigidity is Independence Preserving’ rules out. If the underminer is irrelevant before the perceptual state supports the proposition, it is irrelevant after as well. So Rigidity prevents perceptual undermining when it obviously shouldn’t.” (Weisberg, 2014, p. 126)

---

4See Weisberg (2014, p. 126). For the Classical versions of the Rigidity and RIP principles take the partition to consist of a single cell weighted to 1.

5One could model the undermining effect of *I’m on color-drugs on the bike is green* by updating instead on something like *it appears as if the bike is green*, which in turn raises the probability of *the bike is green* given a low probability in *I’m on color-drugs*, and then regard anything that raises that last probability as an undermining defeater for the greenness of the bike; in the language of Pryor (2013) this could be modeled as a case of ‘quotidian’ undermining. This is beside the point. Weisberg’s puzzle presents a problem for anyone who thinks that the propositions conditionalized upon — whatever those happen to be — can themselves be undermined, and so updating upon beliefs about how things seem offers a solution only if we agree that (i) those beliefs cannot be undermined, and (ii) all cases of undermining defeat are quotidian. An evaluation of that approach is outside of the scope of this essay – I’ll be arguing that the Bayesian has a solution to Weisberg’s puzzle requiring neither indefeasible updates nor pan-quotidianism about undermining defeat.
3 Bayesian Learning More Carefully

Below I’ll be arguing that Weisberg’s conclusion is too strong, but in order to do so we must first take a closer look at Bayesian perceptual learning and the ways that it’s constrained by Rigidity.

3.1 Bayesianism is Incomplete

Bayesianism is at best an incomplete theory of epistemology, in the sense that there are at least two very important varieties of constraints upon rational credence assignments that it is unable to explain. The first variety of incompleteness concerns rationally permissible starting credence functions, functions adopted by agents who possess no evidence at all (so-called ‘super-babies’). There are many intuitively impermissible starting credence functions that are nonetheless perfectly consistent with Probabilism, and hence whose impermissibility cannot be explained by anything within the Bayesian formalism. An agent’s choice of starting credence function will of course determine how perceptually acquired information is to affect other credences via conditionalization, as it will determine their conditional probabilities.

I’ll return to the significance of the Bayesian formalism’s underdetermination of rationally permissible starting credence functions in §5, but right now I want to focus on another type of incompleteness in the Bayesian account of perceptual learning. Just as Probabilism alone is too weak to rule out all of the intuitively impermissible starting credence functions, conditionalization is too weak to rule out all intuitively impermissible credence revisions. That’s because not all permissible credence revisions proceed via conditionalization, and those that don’t are only minimally constrained by the Bayesian formalism.

The most important credence revisions that don’t proceed via conditionalization come as a result of a perceptual experience. Why am I rational in believing that the stove is warm? Because it feels warm. Why am I rational in believing that the cat is on the mat? Because I had a perceptual experience as of a cat on the mat. Those experiences make it rationally per-
missible to form those beliefs. On the sort of subjective Bayesian picture that we’re considering, probabilities are understood as partial belief states, and the only sorts of things that can be partially believed are propositions. Experiences as of warm stoves or cats on mats might have propositional content (I think that they do), but they are not themselves propositions and so they cannot be assigned credences. Hence they are not the sorts of things that can be conditionalized upon. Hence conditionalization cannot be the whole story when it comes to rationally permissible credence revisions.

Bayesians construct formal models of rationally permissible credence revisions, and all of the revisions that they model proceed via conditionalization. As we’ve seen, many rationally permissible credence revisions do not proceed via conditionalization and hence not all rationally permissible credence revisions are modeled. This distinction will be important to what follows, so I’ll introduce some terminology: call the revisions modeled by the Bayesian endogenous credence revisions (as in endogenous to the model), and call the rest exogenous revisions.

3.2 Rigidity and Independence, Carefully This Time

With the distinction between endogenous and exogenous revisions in mind, let’s take a closer look at the rigidity of conditionalization. To that end (and I swear this is relevant) note that it’s common for a single perceptual experience to affect one’s rational confidence in several different propositions. For example, if I have a perceptual experience as of a red, spherical ball, I might shift my confidence in the ball is red and the ball is spherical, along with lots of other propositions (experience is pretty rich, after all). Though

---

6For those who prefer a picture on which agents update on propositions about how things seem rather than how things are, the question becomes: why am I rational in believing I’ve had an experience as of a cat on the mat? The answer is the same: because of my experience.

7Note how minimal I’ve been in describing the role of experience in fostering rationally permissible credence revision. The point applies not only to those (such as myself) who think that a perceptual experience can be evidence that justifies belief, but also to those who think that it can play only a non-evidential, non-justificatory role in making certain beliefs or credence revisions rationally permissible (e.g. Davidson, Jeffrey, and Williamson).

8See [BLIND]; the terminology originates with Howson and Urbach (1993, p. 82).
the details of how to model this phenomenon will differ slightly on the Jeffrey and the Classical Bayesian accounts, they share some important similarities, and in both cases those details have important implications for the rigidity of Bayesian perceptual learning.

Consider first how the Classical Bayesian will model a case in which an agent exogenously revises her credence in more than one proposition at a time. At \( t_1 \) Clara accepts a credence function such that \( P_{t_1}(A) = P_{t_1}(B) = P_{t_1}(A|B) = 0.5 \), and then at \( t_2 \) she exogenously shifts her credences in \( A \) to 1 and in \( B \) to 1. As discussed above, this exogenous shift alone will fix some subset of the probabilities that she accepts at \( t_2 \) — in this case that set will include the probabilities of \( A \) and \( B \) — with others being determined by conditionalizing upon that subset. But what exactly does it mean to update not on a single proposition, but on a set of propositions? For the Classical Bayesian, the answer is very simple: update on all of the new evidence acquired by updating on the conjunction of all of the propositions whose probabilities have just been exogenously revised to 1, which in this case means updating on \( A \cap B \).

We’ve seen that Classical conditionalization is rigid, meaning that updating on \( A \cap B \) never changes the probability of any other proposition conditional on \( A \cap B \). Importantly, though, conditionalizing on that conjunction will not in every case preserve the probability of some proposition \( C \) conditional on one of the conjuncts, \( A \) or \( B \). Suppose for reductio that that’s false, and so for any \( A \), \( B \) and \( C \), \( P(C|A) = P_{A \cap B}(C|A) \). No matter what values are assigned to \( P(C|A) \) and \( P(C|A \cap B) \), it must be the case that \( P_{A \cap B}(C|A) = P_{A \cap B}(C|A \cap B) \); after all, at that point I’ve assigned a credence of 1 both to \( A \) and to \( A \cap B \). The rigidity of conditionalization ensures that \( P_{A \cap B}(C|A \cap B) = P(C|A \cap B) \), and so given our supposition it follows that \( P(C|A) = P(C|A \cap B) \) for any \( A \), \( B \) and \( C \). But this last equality is often false — my credence that the table is delicate given that it’s made out of glass is much higher than my credence that it’s delicate given that it’s made out of glass and it’s incredibly sturdy — and so our supposition is false.

The lesson so far isn’t that Classical conditionalization isn’t rigid; it is.
The two-part lesson is that (i) the proposition that’s conditionalized upon might be just one of the many propositions that are exogenously revised, and (ii) though Classical conditionalization is rigid with respect to the one proposition that’s conditionalized upon, it’s not rigid with respect to those other exogenously revised propositions.

Having appreciated both (i) and (ii) we’re now in a position to sketch a possible response to Weisberg’s puzzle. As far as that puzzle goes, Rigidity is only interesting because rigid updates preserve independence between the proposition updated upon and other propositions. This is puzzling only if we assume that the propositions losing their independence with potential undercuts are the ones that we update upon directly, rather than conjuncts in a larger conjunctive proposition that we update upon. If we drop that assumption then we are free to concede the rigidity of conditionalization without thereby conceding that conditionalization preserves the independence of exogenously revised, perceptually justified propositions with their potential undercuts.

Below I’ll develop this response on behalf of the Jeffrey conditionalizer, but first let’s note that it’s hopeless for the Classical conditionalizer. The case here is overdetermined, but I’ll mention just one reason that’s particularly salient to our discussion. Classically conditionalizing upon $A \& B$ requires assigning a credence of 1, which requires assigning $A$ a credence of 1. But any proposition assigned a credence of 1 is probabilistically independent of any other proposition⁹, and so any proposition that’s been updated upon, or any of their logical implications, will be independent of all other propositions. It follows that if $A$ and $C$ are independent before I Classically conditionalize on $A \& B$ then they’ll be independent afterward. Hence while Classically conditionalizing on $A \& B$ can change credences conditional on $A$ or on $B$, it can’t destroy the independence of $A$ or $B$ with some other proposition.

Jeffrey conditionalization avoids this particular problem by allowing updates on propositions with credences less than 1. For example, Jeffrey Bayesianism describes how to conditionalize when an experience makes it

---

⁹Assuming that the propositions in question have credences greater than zero.
rationally permissible to exogenously revise my credence in the ball is red to .7 and my credence in the ball is spherical to .9. But this creates a new problem: while the Classical Bayesian can handle cases of multiple propositions whose credences have been revised exogenously by condition- alizing on their conjunction, this move in most cases is unavailable to the Jeffrey Bayesian. A probability assignment of 1 to each conjunct ensures that the probability of the conjunction will be 1 as well, but assignments of probabilities between 0 and 1 to the conjuncts is consistent with a range of probability assignments to their conjunction. For example, if I think that $P(\text{the ball is red}) = .7$ and $P(\text{the ball is spherical}) = .9$, the value of $P(\text{the ball is red } \& \text{ the ball is spherical})$ can anywhere between .6 and .7, and where in that interval the probability of that conjunction lies is unde- termined by the probabilities of the conjuncts themselves.

Jeffrey conditionalizers face a second complication in selecting what to conditionalize upon. While Classical Bayesians update on a weighted proposition, Jeffrey Bayesians update on a weighted partition of the state space, where a partition is simply a division of that space into mutually exclusive and jointly exhaustive parts, each weighted according to its probability. Propositions such as the ball is red and the ball is spherical are neither ex- clusive nor exhaustive, and so they typically won’t partition the relevant state space (though they might: see fn. 11).

Jeffrey (1983, p. 173) resolves the issue in a very simple way. He begins by identifying an initial set of propositions — he calls them ‘originating propositions’ — whose probabilities shift exogenously, but which typically are not the $B_i$’s in partition $\{B_i\}$. Those $B_i$’s are instead conjunctions con- structed by taking each originating proposition or its negation as a conjunct. For example, taking $A$ and $B$ as our originating propositions we wind up with four conjunctions as our $B_i$’s: $A \& B$, $A \& \neg B$, $\neg A \& B$ and $\neg A \& \neg B$. These four conjunctions (Jeffrey calls them ‘atoms’; I’ll follow more recent authors and call them ‘elements’) are mutually exclusive and jointly exhaust any probability space, and so taking each conjunction as one of our $B_i$’s, the set $\{B_i\}$ will form a partition, allowing us to Jeffrey conditionalize upon

\[\text{This is slightly misleading — see fn. 11.}\]
it.\textsuperscript{11}

With all of this in mind, let’s revisit Rigidity with an eye to clarifying precisely what’s rigid with respect to what on the Jeffrey picture. Recall that Rigidity says:

\textbf{Jeffrey Rigidity} If $P_{\text{new}}$ is the credence function resulting from accepting $P_{\text{old}}$ and then updating on a shift in partition $\{B_i\}$, then for any proposition $A$ and any $B_i \in \{B_i\}$, $P_{\text{new}}(A|B_i) = P_{\text{old}}(A|B_i)$

We’ve just seen that each $B_i \in \{B_i\}$ is a conjunction of originating propositions, not an originating proposition itself. Hence what Jeffrey Rigidity rules out is a change in conditional probabilities on conjunctions of originating propositions, not changes in conditional probabilities on the originating propositions themselves. As with Classical conditionalization, the probabilities of the individual conjuncts — the originating propositions — conditional on other propositions will not be so-constrained.

Consider again my perceptual experience as of the red, spherical ball. Before that experience I assigned a probability of .5 to each proposition, and I assign $P_{\text{old}}(\text{it’s red} | \text{it’s spherical}) = .5$. Upon having that experience I set $P_{\text{new}}(\text{it’s red})$ to .7 and $P_{\text{new}}(\text{it’s spherical})$ to .9. Since those originating

\textsuperscript{11}Many authors — Weisberg and Pryor included — omit this aspect of Jeffrey’s theory in their summaries. I speculate that this is because in certain circumstances the effect of updating on the originating propositions and updating on the elements is the same, and because Jeffrey’s most widely discussed example of how his system works just happens to be one of those circumstances. In the example we are asked to imagine seeing a cloth in poor lighting, which results results in exogenous revisions to the probabilities of the cloth is green ($=G$), the cloth is blue ($=B$) and the cloth is violet ($=V$). Strictly speaking, this should lead to an update on a partition whose elements include the 8 ($= 2^3$) conjunctions that we can construct from those three originating propositions, yet Jeffrey (together with many later authors discussing this example) omits discussion of the conjunctions and simply talks of updating on these three propositions. The reason that this isn’t disastrous in the current case is because we’re asked to also suppose that the agent seeing the cloth is already certain that nothing is more than one color (all over, at the same time, etc) and is also certain that the cloth is either green, blue or violet. Given those suppositions the probability of five of our eight conjunctions is zero, and so they can safely be ignored as elements of the partition. The three remaining conjunctions will each be closely identified with one of our originating propositions: the cloth is green with $G \land \neg B \land \neg V$, the cloth is blue with $G \land B \land \neg V$, and the cloth is violet with $G \land \neg B \land V$. Given the particulars of the case it’s harmless to speak of updating on a partition with elements $G$, $B$, and $V$, but since those particulars will not generally obtain this harmlessness does not generalize.
propositions don’t form a partition, I now need to assign credences to the four relevant conjunctions. Suppose that I do so as follows:

\begin{align*}
P_{\text{new}}(\text{it’s red } \& \text{ it’s spherical}) &= .6 \\
P_{\text{new}}(\text{it’s red } \& \neg(\text{it’s spherical})) &= .1 \\
P_{\text{new}}(\neg(\text{it’s red }) \& \text{ it’s spherical}) &= .3 \\
P_{\text{new}}(\neg(\text{it’s red}) \& \neg(\text{it’s spherical})) &= 0
\end{align*}

Now my credence in $P_{\text{new}}(\text{it’s red } | \text{ it’s spherical}) = 2/3$. We therefore have a case analogous to the one observed above: we have an episode of perceptual learning in which the probability of an originating proposition on something else has changed, all while respecting the rigidity of Jeffrey conditionalization.

That’s the first lesson of this example. The second lesson is actually a bit more interesting. The exogenously revised values that I assigned to my two originating propositions constrain the values that I assign to the elements of my partition — to my four conjunctions — but do not determine them completely. Since (i) the probability of it’s red conditional on it’s spherical is by definition (we are supposing) the ratio of their conjunction to the unconditional probability of it’s red, and (ii) the probabilities of two propositions underdetermines the probability of their conjunction, it follows that (iii) assigning probabilities to two originating propositions (sometimes) underdetermines the probability of one of them conditional on the other. For example, I might just as easily have assigned the following credences after my observation of the red, spherical ball:

\begin{align*}
P_{\text{new}}(\neg(\text{it’s red}) \& \text{ it’s spherical}) &= .7 \\
P_{\text{new}}(\neg(\text{it’s red}) \& \neg(\text{it’s spherical})) &= 0 \\
P_{\text{new}}(\neg(\text{it’s red}) \& \text{ it’s spherical}) &= .2 \\
P_{\text{new}}(\neg(\text{it’s red}) \& \neg(\text{it’s spherical})) &= .1
\end{align*}

In that case my credence in $P_{\text{new}}(\text{it’s red } | \text{ it’s spherical}) = 7/9$, yet just as before my credences in it’s red and it’s spherical are .7 and .9, respectively.
The probabilities assigned exogenously to originating propositions are only minimally constrained by the Bayesian formalism. As we now see, even once those probabilities are selected the probability of their conjunction is underdetermined, and hence the probability of one originating proposition conditional upon another is also underdetermined. It’s frequently the case that experience makes it rationally permissible to revise the probabilities of several originating propositions at once, and as a result it’s frequently necessary for agents to go further and determine the values of their conjunctions in order to form the partition required for updating. The upshot, then, is that perceptual learning as understood by the Jeffrey Bayesian effectively involves changing conditional probabilities that are unmediated by Jeffrey conditionalization and hence are unconstrained by Rigidity.

We’re now in a position to draw a broader lesson regarding the significance of the rigidity of conditionalization for Bayesian perceptual learning. Episodes of perceptual learning involve an exogenous assignment of credences to some propositions (as a result of an experience or something else) and also an endogenous assignment of credences (via conditionalization) to others. The endogenous assignments reflect the bearing of the exogenously set credences upon the rest.

Weisberg correctly points out that the rigidity of conditionalization prevents the introduction of probabilistic entanglement between the bike is green and I’m on color-drugs via conditionalization. But the intuition driving Weisberg’s Puzzle is not that the probabilistic entanglement is introduced via conditionalization, but merely that it’s one result of my perceptual experience.

As we’ve seen, the Bayesian account of perceptual learning involves more than just conditionalization: it also involves exogenous credence revisions that don’t proceed via conditionalization. Moreover, those exogenous revisions commonly result in changes to the probability of one originating proposition conditional upon another, as we saw in the case of the red spherical ball. Finally, even once the exogenously set unconditional probabilities of
our originating propositions are set, there’s considerable flexibility in setting their new conditional probabilities.

4 Formal Proposal

My proposed response to Weisberg’s puzzle is fairly simple. Intuitively, having a perceptual experience as of the greenness of my daughter’s bike should result in (i) an increase in my confidence in the bike is green, (ii) no change to my confidence in I’m on color-drugs, and (iii) the introduction of a negative correlation between I’m on color-drugs and the bike is green, i.e., it should now be the case that $P_{\text{new}}(\text{the bike is green} \mid \text{I’m on color-drugs}) < P_{\text{new}}(\text{the bike is green})$. Rigidity prevents the introduction of this negative correlation endogenously via conditionalization on a partition that includes the bike is green as an element, and so it must not be an element of the partition. Assuming that my confidence in that proposition is to be increased exogenously, the introduction of the negative correlation in (iii) requires that my new credence in both the bike is green and I’m on color-drugs must be among the conjuncts of the elements of the input partition. Hence what must happen is that my credence in both of those propositions must be set exogenously.

I’ll defend this proposal in §5, but for now let’s just get a sense for how it works out formally. We take as our originating propositions the bike is green (=G) and I’m on color-drugs (=D), and hence we partition our state space into four elements, correlating with the four possible combinations of those propositions and their negations. For simplicity assume that each of the four elements starts out with a probability of 1/4. The introduction of the negative correlations looks like this:
Here my confidence that I was on color drugs when I had my perceptual experience as of the green bike hasn’t changed, and my confidence that the bike is green has increased. If we suppose that I’m on color-drugs is a complete undermining defeater — a defeater that deprives the perceptual experience of all of its evidential force — then if I were to become certain that I was on color drugs then my epistemic situation vis-à-vis the bike is green before I had the perceptual experience should be the same as my situation after having the experience and becoming certain of the underminer. In pictures:

**Figure 1: introduction of negative correlation between D (= I’m on color drugs) and G (= the bike is green).**

If at \( t_3 \) I become more confident that I was on color drugs without becoming certain of it the result is a net decrease in the \( G \) space:

**Figure 2: becoming certain of a full undermining defeater.**
What if instead of varying my degree of confidence in a full undermining defeater, we vary the degree to which $G$ is undermined? For example, suppose that the color-drugs only somewhat decrease the reliability of my color perception, so that $D$ is a \textit{partial} undermining defeater. This will be set at that initial exogenous revision in response to the experience. The particular mechanism will be that it will increase the size of the $D&G$ space at the expense of the $\neg D&G$ space, where a greater increase means a weaker undermining effect. Assuming that this doesn’t reduce my new (at $t_2$) credence of $G$, that means that the ratio of $\neg D&G$ to $\neg D&\neg G$ will decrease slightly. If my confidence in $D$ increases at $t_3$ (but not quite to 1) the picture looks like this:

\begin{figure}[h]
\centering
\begin{tabular}{|c|c|}
\hline
$P_{t_1}$ & $P_{t_2}$ & $P_{t_3}$ \\
\hline
$D&G$ & $\neg D&G$ & $\neg D&G$ \\
$D&\neg G$ & $\neg D&G$ & $\neg D&G$ \\
\hline
\end{tabular}
\caption{Uncertainty of a partial undermining defeater.}
\end{figure}
5 Defending the Proposal

Weisberg’s puzzle illustrates that the introduction of negative probabilistic correlation between an originating propositions and its potential undermining defeater cannot be modeled within the Bayesian formalism. Weisberg takes this to be a reason to reject Bayesianism. I have proposed instead that it is a reason to move the introduction of that correlation outside of the model, so that it is already achieved once conditionalization takes place. I’ve shown that this is consistent with Jeffrey Bayesianism, which already assumed the existence of credence revisions taking place off-model (exogenous revisions) that can change conditional probabilities on propositions involved in those off-model revisions, and so includes a formal mechanism for incorporating those revisions into the model.

Weisberg (2014, p. 142-5) anticipates this type of response, calling it the ‘appeal to richer inputs’, and raises two objections. His first objection is that my proposal requires input partitions that are far more complex than the simple four or eight cell partitions that I’ve diagramed in §4. After all, there are lots and lots of potential undercutters for instances of perceptual learning, and each of them will need to become negatively correlated with the proposition that they have the potential to defeat. On my proposal each of those propositions will need to be treated as an originating proposition, and as a result the input partitions will be fairly fine-grained. Moreover, the determination of which fine-grained partition to adopt given a particular experience will take place outside of the formal model. Hence my proposal involves a loss of explanatory power for Bayesianism. I’ll return to this objection in just a minute.

More troubling to Weisberg than a mere loss of explanatory power is exactly what is being left unexplained:

An update rule is supposed to determine our new credences as a function of our old beliefs and the new evidence. But on the current proposal, “the new evidence” is not really the new evidence. The complex distribution we would be plugging into Jeffrey conditionalization would be produced by considering how an
experience as of a red-looking sock and our background beliefs about optics combine to warrant new beliefs about the quality of the air and the colour of the sock. And this is precisely the kind of work our update rule was supposed to do. (144)

This second objection to my proposal amounts to (i) a proposed criterion of adequacy for any update rule, and (ii) the claim that if my proposal is accepted then conditionalization doesn’t satisfy this proposed criterion. I’ll just grant (ii), and instead argue that the criterion in (i) is not one that conditionalization needs to satisfy. I’ll be assuming that Weisberg’s objection is aimed at my proposal in particular, rather than at Bayesian conditionalization as such; if conditionalization faces insurmountable problems even without my proposal then conditionalization together with my proposal is a non-starter, in which case Weisberg’s objections are superfluous. I show that this assumption is inconsistent with Weisberg’s criterion for an adequate update rule, as that criterion isn’t satisfied by any version of Bayesian conditionalization. I then argue that my proposed constraint on exogenous credence revision is very much in the spirit of other constraints that Bayesians must already accept, and as a result the most reasonable criterion of adequacy for an update rule that’s satisfied by Bayesian conditionalization is also satisfied by my proposed constraint upon that rule.

The first problem with Weisberg’s criterion is that the ‘kind of work’ that’s being demanded of conditionalization is one that no Bayesian update rule — Classical or Jeffrey, with my proposal or without — is capable of doing. Like Weisberg, I think that perceptual experience is one type of evidence. The problem is that perceptual experience is the wrong sort of thing to be conditionalizing upon. As discussed in §3.1, only endogenous credence revisions are governed by a Bayesian update rule, and the only thing that can spark an endogenous revision is an exogenously revised credence. Having a perceptual experience and exogenously revising a credence are two very different things,¹² and hence we never conditionalize upon our experiences, and hence we never conditionalize upon perceptual evidence.

¹²See(Plantinga, 1993, p. 82-3).
It follows that if the adequacy of an update rule demands that it take us from old beliefs and new evidence to a new credence function then Bayesian conditionalization is not an adequate update rule.

Some authors deny that experience can serve as evidence. Taking inspiration from (Davidson, 1986, p. 311), Richard Jeffrey (1983, p. 184-5, 211) holds that only a belief can justify a belief (i.e. can be evidence), and since experiences aren’t beliefs it follows that experiences can’t be evidence. On his (and Davidson’s) view, experience may cause credences to shift, but those shifts are inapt for rational evaluation and hence not within the purview of epistemology. Williamson (2000, p. 197-200) thinks something very similar: though only known propositions count as evidence, some propositions are known because (yes, that’s his term) of the agent’s experiences, which are not themselves evidence.

If all evidence is propositional then my objection to Weisberg’s criterion for an update rule is moot. But this response is inconsistent with the spirit of that criterion. The idea seems to be that an update rule should model the epistemic significance of experiences, whether we call those experiences evidence or something else; this is something that conditionalization cannot do. For Davidson and Jeffrey, note that two agents with identical beliefs/credence functions might not be rationally alike, as one might have some beliefs caused by a perceptual experience, and hence capable of justifying other beliefs, while the other has beliefs with some other causal origin. On this view the epistemic difference between the two agents can’t be explained without accounting for the etiology their beliefs, which will require an account of the relationship between propositions and experiences — between propositions and non-propositions — something Bayesians cannot do within their formal model. Hence even for someone with Jeffrey-like views on perceptual justification, conditionalization cannot satisfy the spirit of Weisberg’s criterion. (I don’t mean to suggest that Jeffrey himself ever thought that it could do something like that; he didn’t. I want simply to dispel the notion that adopting Jeffrey’s views on evidence renders conditionalization consistent with Weisberg’s criterion.)

Williamson’s views are a bit more complicated. Whereas for Jeffrey
beliefs caused by experience can be evidence for other propositions (either by acting as evidence or in some other way), Williamson thinks that only knowledge could play that role. If Jeffrey is right then we can hold the initial beliefs fixed while changing the epistemic status of inferred beliefs by changing the etiology of those initial beliefs. But if it’s knowledge that serves this evidential role then that same trick won’t work, at least not given the rest of Williamson’s view. Williamson thinks that evidential relations are objective relations between propositions: input a set of evidence propositions (the ones that the agent knows) into their credence function and out comes the probability that ought be assigned to every other proposition (op. cit. §10.2). Against Baysians he claims that this probability function itself (in contrast to its inputs) is eternal/ insensitive to the beliefs of the agent. On this view it really doesn’t matter why or on what grounds the agent knows evidence proposition A, only that it is known.

Nonetheless, Williamson’s update rule is also inconsistent with the spirit of Weisberg’s criterion. After all, the epistemologist will still want to know why, in virtue of what, particular agents have the evidence that they do in fact have/ know the things that they know non-inferentially. The epistemological significance of experience does not disappear simply because we stop calling it ‘evidence’. (As above, I don’t mean to suggest that Williamson thinks otherwise.)

Weisberg objects to my appeal to richer inputs by pointing out that it’s inconsistent with his adequacy criterion for an update rule, but we’ve just seen that Bayesian conditionalization is inconsistent with that criterion even without my proposal. That amounts to an objection to Bayesianism as such, which contradicts my original assumption that Weisberg is taking aim at my proposal in particular.

We can try to patch things up by restating our criterion for an adequate update rule: it should determine our new credences as a function of our old credences plus an exogenous revision to our credence function.

Unfortunately, Bayesian conditionalization still fails to meet this revised standard, as our old beliefs together with an exogenous revision sometimes fail to determine what our new credences ought to be. One way that this
can happen is if the agent’s prior credence function is incoherent. If I think that \( P_{\text{old}}(P \& Q) = 1.7 \) and then I conditionalize on \( Q \), then \( P_{\text{new}}(P) \geq 1.7 \), in which case \( P_{\text{new}} \) is incoherent. Another way is if the agent starts out with a coherent credence function but then introduces incoherence at the exogenous phase: when I exogenously revise my credence in \( Q \) to -27, that credence is incorporated into my posterior credence function, rendering it incoherent. Hence if we require of our update rule that it determine a new credence function from an old credence function plus an exogenous revision, and if by ‘determine’ we mean something like ‘specify a credence function that it’s rationally permissible to adopt’, then Bayesian conditionalization fails to meet this requirement. Again, what we have is an objection to Bayesianism as such rather that to my proposal in particular, contradicting my initial assumption.

The response in defense of our revised adequacy criterion is pretty obvious: conditionalization isn’t the only constraint that Bayesians impose upon rationality; they also accept Probabilism, the thesis that all rationally permissible credence functions must satisfy the probability axioms. Conditionalization determines a rationally permissible credence function only from a rationally permissible exogenous credence revision together with a rationally permissible prior credence function, and incoherent credence functions or incoherent exogenous revisions aren’t rationally permissible.

Thus we can further revise our account of what conditionalization should do: it should determine a new, rationally permissible credence function for any agent who starts out with a rationally permissible credence function and then makes a rationally permissible exogenous revision.

We’re finally in a position to make the point that I’ve been working up to. If the Bayesian formalism is only capable of explaining transitions from a permissible credence function plus a permissible exogenous revision, then the more constraints that we impose upon the permissibility of such functions and revisions, the less our formalism is capable of explaining. I’ve proposed that we must impose an additional constraint upon the exogenous revisions: they must encode the negative probabilistic correlations associated with potential undermining defeaters for any belief that’s directly affected by an
experience. Weisberg objects that that’s not an appropriate bit of work to be left unexplained by our formalism. But as I argue below, Bayesians are already committed to a closely analogous constraint upon starting credence functions, and it’s unclear why the one constraint should be considered more problematic that the other. Hence it’s unclear why Weisberg’s objection to my proposal doesn’t generalize into a broader indictment of Bayesianism.

As noted, Probabilism ensures that certain evidential relations will be encoded in any permissible credence function. For example, it ensures that any evidence that makes is rationally permissible to set $P_{\text{new}}(A)$ to .7 also makes it rationally permissible to set $P_{\text{new}}(\neg A)$ to .3, and prohibits setting $P_{\text{new}}(\neg A \& B)$ any higher than that. But not all intuitively mandatory evidential relations — those to which all rational agents are obliged to conform — are encoded by Probabilism, and hence many probabilistically coherent credence functions are intuitively impermissible. Famously, Probabilism fails to ensure that the observation of lots of green emeralds and no non-green ones supports ($H_1$) all emeralds are green more than it supports ($H_2$) all emeralds are grue.\textsuperscript{13} Because both $H_1$ and $H_2$ entail $E = \text{all observed emeralds are green}$, conditionalizing on $E$ will increase (or leave the same) my confidence in both of them, and yet intuitively my posterior credence in $H_1$ should be much higher than that of $H_2$. According to the Bayesian, that means that before acquiring evidence $E$ it must be the case that $P_{\text{old}}(H_1 \& E) > P_{\text{old}}(H_2 \& E)$. In other words, if we wish to ensure that conditionalization determines a rationally permissible credence function from an update on $E$ the we must constrain our prior credence functions in ways that go well beyond Probabilism.

In many cases this phenomenon appears innocuous, as when an agent starts out thinking that $P_{\text{start}}(H_1 \& E) \leq P_{\text{start}}(H_2 \& E)$ and then adopts the desired inequality after acquiring new evidence and updating in the normal way. But this merely pushes the bump in the rug. Take $E^*$ to be the conjunction of $E$ and all of the other evidence that the agent has acquired to $t$. In that case a necessary and sufficient condition for the agent holding that $P_t(H_1) > P_t(H_2)$ is that their starting credence function $P_{\text{start}}$ be such

\textsuperscript{13}See Goodman (1946) and (1983, p. 72-83).
that $P_{\text{start}}(H_1 \& E^*) > P_{\text{start}}(H_2 \& E^*)$.\textsuperscript{14}

The narrow point is that if the Bayesian is to regard inductive inference as more epistemically respectable than counter-induction or non-induction, they’ll need to go beyond mere Probabilism and impose further constraints upon starting credence functions. The broader point is this: for any $E$ and $H$ such that $E \not\equiv H$ and $H \not\equiv E$, ensuring that $E$ supports $H$ more than some competing hypothesis depends crucially on the choice of starting credence function. We have lots of intuitions about evidential relations that go beyond deductive entailment (e.g. the intuition that induction is preferable to counter-induction), and in order to require of agents that they satisfy those intuitions we have to constrain their starting credence functions. The Bayesian formalism (= Probabilism + conditionalization) doesn’t impose those restrictions, and hence additional constraints on starting credence functions are needed in order to ensure that their prior credence functions are rationally permissible, which are themselves required in order for conditionalization to determine a rationally permissible posterior credence function given some permissible exogenous revision.

For the Bayesian, there are obvious parallels between what’s required by the grue paradox and what I’m proposing in response to Weisberg’s puzzle: just as the former requires a constraint upon rationally permissible starting credence functions, the latter requires a constraint upon exogenous revisions. Ideally, both of those constraints would be imposed by the formalism itself, but in both cases that’s proven not to be the case. If we assume that Weisberg is objecting to proposals like mine, rather than to Bayesianism in general, then the problem can’t simply be with the existence of intuitively compelling constraints upon the formalism for which we have no widely accepted formal theory; we have no such theory for distinguishing ‘projectable’ predicates like green from ‘unprojectable’ ones like grue, either. Presumably, then, the objection must be either (i) that such constraints are more objectionable at the exogenous revision side of the model than at the starting credence function side, or (ii) to some other feature of undermining defeat

\textsuperscript{14}For Classical Bayesianism at least – the possible non-commutativity of Jeffrey Bayesianism makes that case less straightforward. See Domotor (1980).
that distinguishes it from broader inductive practice, and in virtue of which my proposal is the more problematic. (i) seems arbitrary, and (ii) is not forthcoming. Seen in this light it’s unclear how my proposal presents any special problem for the Bayesian that isn’t closely analogous to a problem that they already have, and hence it’s unclear how Weisberg is objecting to my proposal in particular rather than to Bayesianism in general.

We’re left, then, with Weisberg’s first objection — that my proposal involves a loss of explanatory power — and in this he is completely correct. It is unwelcome news that the Bayesian is unable to model the introduction of a negative correlation between an exogenously revised proposition and its undercutters. He’s also correct that the input partitions will need to be more complicated than those in my examples from §4, as each potential underminer will need to be taken as one of our originating propositions. Furthermore, it’s doubtful that those complicated input partitions will be among the ‘direct effects’ of experience and hence it’s likely that the auxiliary theory bridging the gap between experience and input partition will be more complex than the Bayesian might have initially supposed.

These are real objections to my proposal, and the best that can be done in response is to mitigate their badness. Two considerations to that effect. First, we again find a close parallel in our response to the grue paradox, which itself involves an enormous loss of explanatory power relative to what a Bayesian might have expected; if that’s something that the Bayesian can live with then so is the loss of explanatory power associated with my proposal. Second, though the input partitions required by my proposal will involve a significant number of originating propositions, that number is dwarfed by the number of propositions that are not involved in it. Though the formal model will be unable to explain the introduced negative correlation between perceptually justified beliefs and their potential undercutters, it will be able to explain how those changes ought to affect the agent’s credences in all other propositions and hence to determine a posterior credence function. Even in its reduced state the explanatory power of the Bayesian formalism is quite robust.

There’s one last objection that I’d like to consider. Weisberg (2014,
p. 129) considers a response a bit like mine, which he characterizes as the claim that Jeffrey conditionalization doesn’t ‘apply’ in cases of perpetual undercutting. The thought here seems to be rooted in the late career Richard Jeffrey’s somewhat unorthodox views about the motivations for conditionalizing. One prominent view among Bayesians is that agents ought to conditionalize because failure to do so leads to the sort of pragmatic defeat illustrated by the Lewis/ Teller dynamic Dutch book argument. Teller (1973) Jeffrey thinks that such considerations are beside the point, and conditionalization is motivated — when it is motivated — by considerations of coherence alone. The Total Probability theorem follows from the probability axioms plus the definition of conditional probability:

**Total Probability**  \[ P_{new}(A) = \sum_i P_{new}(A|B_i)P_{new}(B_i) \]

When the transition from an agent’s old credences to her new ones is rigid on some \( B_i \), \( P_{new}(A|B_i) = P_{old}(A|B_i) \). Hence in such cases, by simple substitution on Total Probability we get:

**Jeffrey conditionalization**  \[ P_{new}(A) = \sum_i P_{old}(A|B_i)P_{new}(B_i) \]

The upshot is that concerns of synchronic coherence alone require that we (Jeffrey) Conditionalize upon our new evidence any time Rigidity holds. That just leaves us with the following question: when does Rigidity hold? Not always, says Jeffrey. And therein lies a possible answer to the puzzle: perhaps cases involving undermining defeaters are cases in which Rigidity does not hold, and hence they are cases in which Jeffrey conditionalization is unmotivated and inappropriate.

To this Weisberg very reasonably objects that perceptual justification is nearly always vulnerable to undermining defeat, and hence if Jeffrey condi-

---

15See Wagner (2013) for a defense of this view.
16Interestingly, Jeffrey (2004) motivates Total Probability with a Dutch Book argument (§1.4) and then goes on to motivate Jeffrey conditionalization by appeal to Total Probability (§3.2). Hence there’s a sense in which he does rely on considerations of pragmatic defeat to motivate conditionalization, but only because they motivate Probabilism.
17Total Probability is a constraint upon all coherent probability functions; \( P_{new} \) is just a special case of the general principle.
tionalization is inapplicable in cases involving the possibility of undercutters then it’s inapplicable in nearly every case of perceptual learning.

Weisberg is no doubt correct about the ubiquity of potential undercutters for perceptual experience, and so if conditionalization is to be rejected in all such cases then rational agents won’t be doing much conditionalizing. But while this might be a serious problem for some other proposals, it’s no objection to mine (to be clear: Weisberg never says that it is). As we observed above, no case of perceptual learning is such that it can be explained completely in terms of conditionalization, since conditionalization is insensitive to experiences and hence cannot account for the exogenous credence revisions that provide the inputs to the conditionalization process. In other words, exogenous revisions that do not proceed via conditionalization are a necessary component of Bayesian perceptual learning. But it doesn’t follow that we must for that reason abandon conditionalization in all cases of perceptual learning. If a coherent agent makes an exogenous revision to some proper subset of their credences and fails to conditionalize then they’re likely to wind up incoherent. Hence unless they respond to an experience by adopting an entirely new credence function they will need to conditionalize or find some other mechanism for maintaining their coherence.

On my proposal, conditionalization applies in every case of perceptual learning, it just doesn’t explain all of the credence revisions that result from perceptual experience. In this regard it’s just like every other version of Bayesianism.

6 Conclusion

The introduction of a negative correlation is an essential aspect of acquiring new information that is itself vulnerable to undermining defeat. Weisberg’s puzzle is important because it illustrates that Bayesians can’t model this effect in any straightforward way. Weisberg himself concludes that this is a reason to reject subjective Bayesianism. I have argued that this conclusion is too strong – the lesson instead is that Bayesians should reduce their explanatory ambitions, moving problematic aspects of undermining defeat
off-model. This move is appealing for several reasons. First, it solves the problem by restoring the consistency of the Bayesian formalism with our intuitions about undermining defeat. Second, the Bayesian account of perceptual learning has always presupposed that some credences will be revised exogenously, revisions that do not proceed via conditionalization, and so my proposal represents only an incremental increase to an already existing aspect of the Bayesian theory rather than a new, dramatic departure. Third, while conditionalization is rigid, Bayesian perceptual learning is not: since both Classical and Jeffrey Bayesians are committed to exogenous revisions that change the ratio of the probability of conjunctions to the probability of their conjuncts, it’s inevitable that conditional probabilities themselves will change exogenously. So again, what I’m proposing isn’t a great departure. Fourth, my proposal doesn’t involve commitment to any cases in which conditionalization doesn’t ‘apply’ – such cases are of course possible, but my proposal is orthogonal to the frequency of its application. Fifth, and finally, there’s a long tradition of Bayesians imposing extra-formal constraints upon their theory in order to deal with counter-intuitive consequences of the minimal Probabilism + conditionalization account, as they do in response to the grue paradox.

References


