Commutativity or Holism? A Dilemma for Jeffrey Conditionalizers

Jonathan Weisberg

1 Introduction

If something looks red, whether or not you should believe that it is red depends on what you think about the reliability of your vision, the quality of the lighting, whether or not you're under the influence of hallucinogenic drugs, and so on. In general, a belief's empirical justification is sensitive to background belief. Call this general view *confirmational holism*. Respecting confirmational holism looks to be a desideratum on any adequate rule for updating our beliefs in response to empirical data. Another commonly held desideratum is *commutativity*, the view that the order in which information is learned should not matter to the conclusions we ultimately draw, provided the same total information is collected. It shouldn't matter whether I find the murder weapon in the maid's room first and then hear testimony about her alibi, or the other way around. Either way my attitude about her guilt should be one of guarded suspicion.

Commutativity and holism both have strong intuitive pull, and it will be an unhappy state of affairs if our epistemology cannot satisfy both. My aim here is to consider whether Bayesian epistemology can satisfy them both. In particular, I want to explore the worry that Bayesianism's standard updating rules — Strict Conditionalization and it's more general counterpart, Jeffrey Conditionalization — have built in limitations in this regard. Worries about the compatibility of Jeffrey Conditionalization and commutativity are familiar, though recent work seems to have more or less resolved that problem. Whether or not Jeffrey Conditionalization can be appropriately holistic is, however, a question that's received far less attention, and it's been thought not to be a problem when it is considered. But the recent work on commutativity seems to point toward a necessary tension with holism. Thus we are threatened with an unappealing dilemma: if we endorse Jeffrey Conditionalization, we must choose between commutativity and holism. The purpose of this paper is to lay out the argument for this dilemma and propose a diagnosis.

To appreciate the problem we have to look at the work that's been done on commutativity and a related problem I'll be calling *the inputs problem*. My approach will be to start with a fair bit of background, beginning with the basics of Strict and Jeffrey Conditionalization. I'll then proceed by outlining the work that's been done on the inputs problem, holism, and commutativity, before developing that work into an argument that holism and commutativity are incompatible on a JC account. I'll close by proposing a diagnosis of the problem.

2 Background

Consider the classic Bayesian rule for updating, Strict Conditionalization:

SC When you learn a proposition E, change your credences from p to $q = p(\cdot|E)$.

As is often noted, SC entails that learned proposition always get credence 1, which is unacceptable. Rarely, if ever, should we be maximally certain of anything. At the very least, most of our perceptual experience does not license certainty in any proposition, so SC does not tell us how to learn from experience. This problem with SC led Jeffrey (1965) to proffer an alternate rule for updating:

JC When experience directly changes your credences over a partition $\{E_i\}$ from $p(E_i)$ to $q(E_i)$, set your new credences to $q = \sum_i p(\cdot|E_i)q(E_i)$.

Jeffrey Conditionalization solves the latter problem by enlarging the scope of SC to accommodate uncertain learning. Experience may not furnish us with any certain information, but it can give us uncertain information in the form of a distribution over a partition, which can then be assimilated into our corpus of beliefs via JC.

At first blush, JC looks pretty good as a replacement for SC, but it has garnered its share of criticism. Perhaps most notorious is the non-commutativity objection. Applying JC to two sets of input values in different orders yields different final results, but acquiring the same total information ought to yield the same final assessment, no matter in what order the information is acquired. JC has also been accused of being incomplete. JC tells us how to respond to probabilities directly affected by experience, but does not tell us how experience ought to determine those probabilities. What partition should an experience directly affect, and what should the input values be? These questions must be answered for us to have a complete account of rational learning, and JC does not answer them.

2.1 The Inputs Problem

JC tells us what our posteriors should be, given our priors and the posterior values over a partition. This is not quite what we want from a theory of rational learning. What we want is to know what we ought to believe as our subjective state changes from one moment to the next. At the very least, we want a theory about how our beliefs ought to change when we have this or that experience: a function that takes an experience and our priors as arguments and delivers posteriors as output. JC would do this job if supplemented with a rule going from experience and priors to the input values for JC. Jeffrey and Carnap (Jeffrey, 1975) discussed the need for a supplementary rule but did not come up with one, and Jeffrey did not give the problem much attention in writing. Skyrms (1975) also acknowledges the gap but has little to say about it. The problem is actually quite serious though, since, without a solution, JC is actually vacuous. Unless we say what partition is directly affected by an experience and what values it ought to get, we can do any update we want without violating JC. To get an arbitrary q from p via JC, we just do a JC update on the set of possible worlds $\{w_i\}$ with the values $q(w_i)$. In applications, of course, we generally have an intuitive understanding of the rule that rules this sort of thing out. But in many cases intuition fails us, and we need a proper theory.

The first attempt I know of to face up to the inputs problem is Field (1978). Field acknowledges the seriousness of the problem and proposes as a solution that experiences be represented by a numerical parameter α, which captures the extent to which an experience supports the input proposition.[1] Field suggested that α be the bayes factor:

$$\beta_{q,p}(E:\overline{E}) = \frac{q(E)}{q(\overline{E})} / \frac{p(E)}{p(\overline{E})},$$

where p is a prior function and q the appropriate posterior one. Once each experience is assigned an α-value, $q(E)$ can be obtained from an experience-prior pair by fixing $p(E)$ in the above equation and plugging the α-value in for $\beta_{q,p}$. Interestingly, I have not seen it noted that Field's proposal is the barest sketch of a solution since it does not actually say anything about what experiences get which α-values, or even which E should go with a given experience. That may be because Field's proposal, bare as it is, ran into an immediate problem.

Garber (1980) objects that Field's proposal allows repetitions of the same experience to yield unbounded support for the input proposition. If the same experience is always associated with the same α, then updating on the same experience a few times in a row can yield near certainty in E. But, says Garber, the same experience repeated over and over should not be so informative. The source of the problem seems to be that α fixes $q(E)$ as a function $p(E)$ alone, so that the effect of the experience on an agent is the same regardless of her other background beliefs, such as that she just had that same experience. But Garber's objection may turn on mishandling repetitions of the same experience. If I have the same experience at t_0 and then t_1, maybe those experiences should bear on two different partitions, to reflect the change in time. $\mathbf{E}_0$ should contain propositions about what things are

[1] For ease of exposition I'm following Field in sticking to the simple case where the input partition is $\{E, \overline{E}\}$, and we can speak of the input proposition, E. Nothing hangs on this simplification.

like at t_0, and $\mathbf{E}_1$ should be about t_1.[2] So the same experience repeated doesn't amount to ramping up the probabilities on a single proposition.

Nevertheless, α still fixes $q(E)$ as a function of $p(E)$ alone, and we can use this fact to recast Garber's objection. Suppose the experience in question is of a blue-looking cloth, and E is *the cloth is blue at t*. Then, even if the agent thinks there's tricky lighting around, Field's rule requires her to draw the same conclusion about the cloth's color as she would if she thought the lighting were good. Whatever her prior that the cloth is blue, the experience will give that credence the same boost whether or not she has reason to distrust the appearance of things. It seems the input values to JC shouldn't just depend on the prior value of E, so Field's proposal can't be right.[3]

For a while, that was the end of discussion of the issue. Room was left open for alternative proposals but nobody offered one. Christensen (1992) then picked up the question in connection with an apparent problem for JC deriving from holist epistemology. Christensen wonders whether the lesson of holist epistemology — that no belief's empirical justification is entirely independent of the agent's background theory — is bad news for JC. The worry is that JC depends on being able to isolate a proposition whose probability can be fixed based on experience, and from which all other beliefs can be determined. But holism says that every proposition's probability depends on more than just the experience. Christensen concluded (correctly) that the *prima facie* problem is only apparent. Holism says that $q(E)$ doesn't just depend on the experience, or even on just the experience and $p(E)$, but on the experience together with the entirety of p. This requirement does rule out Field's proposal, as Doring (1999) has pointed out. But it doesn't preclude that, once $q(E)$ is fixed, the rest of q cannot be determined from p and $q(E)$, which is all that JC assumes. Christensen concludes that the completion of a JC-based account is not precluded by holism, though the dependence of $q(E)$ on the entirety of p suggests that the project will not be simple.

And that's where things stand on the inputs problem. To provide a complete JC-based account we need a rule describing the dependence of $q(E)$ on p and the experience. From Field, Garber, and Christensen we know that $q(E)$ should depend on the entirety of p, but we have no idea what this dependence will look like. So far as I know, no further work has been done, probably because the project is a daunting one. Surprisingly, otherwise promising work towards a solution of the apparently unrelated commutativity problem suggests that the project may

[2] Even if the partitions more realistically contain temporal indexicals, the change in time will instead be reflected in a change in those propositions' values, so that iterating JC updates still doesn't result in a monotonic increase of E's probability.

[3] You might think that we're still using the wrong E; that E should be something like *the cloth looks blue at t*. Maybe this variant of Garber's problem is always fixable by appeal to appearance propositions. I'll return to that question when I discuss holism and commutativity.

actually be impossible.

2.2 Commutativity

JC is non-commutative on input values. This is most easily seen when we consider subsequent Jeffrey updates on the same partition $\mathbf{E}$ with the distinct vectors $\langle x_i \rangle$ and $\langle y_i \rangle$. The first update leaves E_i with its input value, x_i, and the second leaves it with y_i. Reversing the order of the updates, E_i gets y_i first and then x_i in the end. Even if subsequent JC updates aren't allowed to operate on the same partition, similar problems still arise. So order matters. Jeffrey first noted this feature of his rule but didn't see it as problematic. But there is a strong intuition that the same total information ought to yield the same final epistemic state, regardless of the order in which the information is learned. Since JC is clearly non-commutative on input values, the only way to respect this intuition is to insist that reversing the order of input values does not amount to learning the same information in reverse order. As Lange (2000) points out, far from being an ad hoc move, rejecting the demand for commutativity on input values is intuitive. For example, having the same experiences in different orders won't just yield the same input values in a different order. To illustrate, suppose I am looking at a cloth in dim light. A first glance (E) moves my credence in its blueness from .1 to .8, and a second glance (E') moves me from .8 to point .9. Now suppose those input values are reversed. In that case my first glance must have been much more clearly blue than E' was, and my second glance must have been much less clearly blue than E was. So reversing input values does not correspond to reversing the order of experiences. But the experiences are the source of the information here, so it seems the experiences ought to commute instead of the input values.

Interestingly, Field's proposed α parameter does yield commutativity on experiences. As Field notes, doing Jeffrey updates on two bayes factors yields the same final state regardless of the order of the updates. But, as Garber's objection illustrates, experiences cannot be identified with bayes factors since this makes empirical learning insensitive to background beliefs. This leaves us in the unfortunate position of having one answer to the inputs problem, Field's, that is commutative but not holistic. The worry we're going to consider now is that this unfortunate position is permanent, since Field's answer to the inputs problem may be the only possible answer that satisfies commutativity.

3 The Tension with Holism

For a while, Garber's objection left open the possibility that some other rule for fixing $q(E)$ could be found that (i) made JC commutative on experiences and (ii) was appropriately sensitive to background beliefs. But Wagner (2002) shows that

Field's rule is often necessary for commutativity. Wagner shows that, on certain minimal assumptions, if two updates commute then they must yield the same bayes factor in both cases:

Wagner's Theorem Suppose p is the initial probability function and we are considering two possible sequences of JC updates it might undergo: first to q and then to r, or first to q' and then to r'. Suppose also that the updates happen on the partitions $\mathbf{E} = \{E_i\}$ and $\mathbf{F} = \{F_j\}$ in opposing orders so that, schematically, we have

$$p \xrightarrow{\mathbf{E}} q \xrightarrow{\mathbf{F}} r$$
$$p \xrightarrow{\mathbf{F}} q' \xrightarrow{\mathbf{E}} r'$$

Then, if $r = r'$ and

$$\forall i_1 \forall i_2 \exists j : p(E_{i_1} F_j) p(E_{i_2} F_j) > 0 \tag{1}$$

$$\forall j_1 \forall j_2 \exists i : p(E_i F_{j_1}) p(E_i F_{j_2}) > 0, \tag{2}$$

we are guaranteed the bayes factor identities

$$\forall i \forall j : \beta_{q,p}(E_i : E_j) = \beta_{r',q'}(E_i : E_j) \tag{3}$$

$$\forall i \forall j : \beta_{r,q}(F_i : F_j) = \beta_{q',p}(F_i : F_j). \tag{4}$$

That is, provided p regards any two elements of $\mathbf{E}$ as probabilistically consistent with some element of $\mathbf{F}$ and vice versa, then commutativity implies that the q and r' updates yield the same bayes factors on $\mathbf{E}$, as must the q' and r updates on $\mathbf{F}$. Thus if two experiences are to commute, it looks like the input values they yield must be insensitive to background belief.

What should we make of Wagner's result? Rather than conclude that JC is faulty, we might reject experience commutativity. After all, reversing the order of experience shouldn't always lead to the same conclusions; watching a window close suggests very different things from watching it open. Still, the order of *some* experiences doesn't matter to *some* hypotheses. Whether I see or hear the rain first, I'll still conclude that it's raining with the same level of certainty. And Wagner's proof can be tweaked to show that this sort of limited commutativity still entails bayes factor identity. Could we, instead, object that Wagner's result does not properly account for the passage of time? Wagner considers the case where a distribution is updated twice on the basis of experience, which doesn't take account of the agent's awareness that time is passing. But again, a few tweaks to Wagner's result show that, in the relevant sort of case, temporal factors won't make a difference: bayes factor identity is still required.[4]

[4]See the appendix for an appropriately tweaked reconstruction of Wagner's result.

It looks like we have to accept at least some anti-holism with Jeffrey Conditionalization. Still, we aren't forced into a thoroughgoing rejection of holism. Notice, for example, that we aren't forced into Field's account, where the same experience must always yield the same bayes factor. For one thing, Wagner's result only applies when (1) and (2) hold, which they don't when $\mathbf{E} = \mathbf{F}$ for example. Of course that's not much comfort if we assume that the partitions affected by subsequent updates are about different times and, hence, never identical. Still, Wagner's result only tells us that an experience **e** must be assigned the same bayes factor across a range of probability functions: those that can be obtained from p by updating on an experience **f** that ought to commute with **e**.[5] Assuming that not every credence function is obtainable from such an **f**, **e**'s bayes factor can vary with background belief to at least some extent.

The question is whether **e**'s bayes factor can vary enough to get the right results. Can there be an experience **f** that ought to commute with **e** but which changes the background assumptions in such a way that the import of **e** should be affected? We are looking for a case where an experience **e** prompts an update on **E**, but the bearing it has on **E** should be affected by the change **f** effects on **F**. A simple example would be a case where $\mathbf{E} = \{E, \overline{E}\}$ and $\mathbf{F} = \{F, \overline{F}\}$, and the first experience is a reason for believing E while the second experience is a reason for believing F, where F is a defeater for the $\mathbf{e} \rightarrow E$ support relation. So suppose **e** is the appearance of a blue cloth, and **f** is the appearance of blue-tinted light fixtures. Ordinarily a visual experience directly warrants the belief that things are as they appear, so let **E** be *the cloth is blue*. Supposing that **e** is had first, the agent should boost her confidence in E significantly, which means that $\beta_{q,p}(E : \overline{E})$ will be something quite large. But if commutativity is to be respected, (3) then demands that $\beta_{r',q'}(E : \overline{E})$ be quite large as well. Thus, even if the agent has experience **f** first and becomes confident that the lighting is blue-tinted, she will still have to become confident that the cloth is blue on the grounds that it appears to be. In short, respecting commutativity requires that **e**'s support for E be insensitive to the agent's background belief in F.

This looks pretty bad, but there's an obvious response: that we have applied JC to the wrong partition. **e** doesn't bear directly on the proposition that the cloth is blue, **e**'s bearing on E is mediated by the fact that it looks like the cloth is blue. So we ought to have E be *It looks like the cloth is blue*. Then we can have $\beta_{q,p}(E : \overline{E}) = \beta_{r',q'}(E : \overline{E})$ be high, since **f** no longer defeats the $\mathbf{e} \rightarrow E$ connection. But for this strategy to work in general, there must always be this sort of mediating proposition available. That is, we need to always have an appearance proposition E available such that $q(E)$ is a function of just **e** and $p(E)$, and such that the rest of q is a function of $q(E)$. Essentially, we are considering rejecting holism and

[5] More precisely: an experience **f** that ought to commute with **e** at least as far as the final probabilities of $E_i F_j$ conjunctions are concerned.

embracing a sort of foundationalist Bayesianism, where appearance propositions serve as the foundation.

Should Bayesians worry about being forced into this sort of foundationalist view? Christensen worries that such a view is not tenable because there cannot be the sort of uncertain, epistemically basic beliefs the view needs. According to Christensen, the very fact that those beliefs would be uncertain betrays their dependence on considerations of background belief. When applying Jeffrey Conditionalization, why must the input values over $\mathbf{E}$ be uncertain? According to Christensen, because $\mathbf{e}$'s support for E can be undercut by background beliefs about the reliability of the $\mathbf{e} \rightarrow E$ connection in the given scenario.

Christensen is certainly right that an experience's bearing on this or that fact can be sensitive to background beliefs, but there might be other reasons why an experience's bearing is uncertain. Consider the way we learn to identify sounds and colors. I can't describe the features of an aural experience that allow me to identify it as my favorite singer's voice, nor can I describe the details of a visual experience that allow me to recognize a particular color. Recognizing colors and sounds is an uncertain process, but not clearly because of any background beliefs we have. Maybe it's uncertain because it is an imperfect, habit-based skill. Arguably, my unconscious identification of phenomenal appearances would be just as (un)certain as it is no matter what I thought about my situation. In fact, Oved and Pollock's (2005) view of the transition from visual experience to belief endorses just this picture. On their view, a visual experience has a *look* and an *appearance*. The look is the raw visual data at the last stage of pre-conscious processing, and we unconsciously track features of this visual data to get its appearance, which is then available for conscious consideration. The appearance contains concepts like color, but the look does not. Colors have to be identified based on the look, and this is an automatic reaction trained through experience. In general, Oved and Pollock think that a feature P is identified based on its look by a trained widget called a P-detector. But the features the P-detector tracks do not make it up to the level of conscious consideration, and the widget operates independently of background belief. I don't know whether this view is actually right, but that's not the important thing here. The point is that there are viable views on which experience would furnish us with uncertain beliefs that are not mediated by our other background beliefs. If P-detectors operate independently of our background beliefs, but yield imperfect judgments, then it makes sense to have the epistemically basic yet uncertain beliefs needed by JC.

This point should give some comfort to Bayesians who were worried by Christensen's argument against uncertain epistemic foundations. But it doesn't get JC out of the woods. After all, if we can always find a background belief that defeats the support an E-looking experience gives to the belief that E, couldn't we play the same game with the proposition that things look E (call it E^*)? If we let $\mathbf{f}$

be a sighting of George, the mad scientist who always leads us to form mistaken beliefs about what color experiences we're having, then the same sort of problem case arises. The visual experience of a blue cloth, if had before George is sighted, should make me very confident in E^*. But if commutativity is to be respected, (3) tells me I have to be confident that that is the visual experience I'm having even if I've already spotted George. Using appearance propositions as an epistemic foundation just denies the thoroughness of holism. If we are taking holism seriously, this is not an acceptable state of affairs.

4 Diagnosis

It's both discouraging and puzzling that something as sensible as Jeffrey's rule should force a choice between such plausible considerations as holism and commutativity. Even if we can't avoid the choice, it seems we should be able to develop at least some appreciation for why it is being forced upon us. Where does JC go wrong? What assumptions do we make in adopting the rule that force us to choose between holism and commutativity? The answer, I want to suggest, actually has less to do with commutativity and more to do with the inexpressibility of a special kind of evidential relation in a JC-based account. Specifically: JC does not allow for the doxastic defeat of a non-doxastic reason.

Let me explain. A doxastic reason for believing A is a belief in another proposition B that is evidence for A. A non-doxastic reason is some other cognitive state, like a perceptual experience, that supports the belief in A without itself being a belief. Bayesianism in general, and JC in particular, express doxastic reasons through conditional probabilities. Roughly, A is a doxastic reason for B if $p(B|A) > p(B)$. But non-doxastic reasons don't get any direct expression in the typical Bayesian apparatus. JC introduces their expression by allowing for input values to be 'directly' affected by experience. The problem is that JC does not allow for *defeaters* of non-doxastic reasons. Bayesianism expresses defeat of a doxastic reason via conditional probabilities — A is a reason for B with C as a defeater if $p(B|A) > p(B)$ but $p(B|AC) = p(B)$. But how could JC capture the fact that a belief is a defeater for the support an *experience* gives to a proposition? How do we capture the idea that F is a defeater for $\mathbf{e}$ as a non-doxastic reason for E? We would need the discovery that F to reduce the probability of E to its initial state if it was raised via $\mathbf{e}$ but not otherwise. The simple, stubborn fact is that JC does not allow this. On a JC-based account, if F is going to reduce the probability of E, it's because F is a reason against E, not because F undercuts the reason that supported E.

To see this, suppose that $\mathbf{e}$ has boosted the probability of E, and that F is a defeater for that support relation. Then F must be prepared to reduce E's probability back to what it was, so $q(E|F) < q(E)$. It follows that $q(F|E) < q(F)$, and so $p(F|E) < q(F)$ since JC is rigid. Because $q(F)$ was determined by JC on $\mathbf{E}$, we then have that $p(F|\overline{E}) > p(F)$, which entails that $p(\overline{E}|F) > p(\overline{E})$. That is, F is a reason for $\overline{E}$, which is equivalent to its being a reason against E. The bottom line is that F can't be a defeater for the $\mathbf{e} \rightarrow E$ support relation after the fact without just being a evidence against E. Succinctly put, JC doesn't allow for after-the-fact defeaters of non-doxastic reasons.

This in itself amounts to a partial rejection of holism, since it entails that $\mathbf{e}$'s initial support for E can't be undercut by a defeater that is discovered later. The role commutativity plays is in making this partial rejection of holism total, by making it time-symmetric. Demanding commutativity ensures that $\mathbf{e}$'s support for E won't be undercut by the defeater even if the defeater is discovered first. Since F can't undercut the $\mathbf{e} \rightarrow E$ support relation after the fact, E will remain supported after F is discovered. If we demand commutativity, we ensure that $\mathbf{e}$ supports E even after F is discovered, which is to say that $\mathbf{e}$'s import is thoroughly insensitive to background information F.

On reflection then, it looks like the problem we've uncovered isn't really a problem with commutativity, but a problem with the expression of non-doxastic defeat in a JC-based account. Wagner's commutativity result just highlights the fact that non-doxastic defeaters are ignored by giving the fact definite shape in a particular application. The problem we face is really that JC doesn't allow for defeat of non-doxastic reasons. The grim conclusion, it seems, is that JC is inherently non-holistic.

There is a vast literature on foundationalist epistemologies, and I don't want to pretend that we can settle questions about the tenability of anti-holism here. The point I want to make is just the more modest one about JC's commitment to this sort of anti-holism. Christensen's concern that proponents of JC can't reject holism because uncertainty is incompatible with epistemic basicness is faulty, as shown by theories like Oved and Pollock's. But simple Quinean considerations still give holism its kick: for just about any support relation we can think of, we can find a background belief that would defeat it. Since the holism/foundationalism question is unsettled this may not spell certain disaster for JC-based accounts. But it does reduce their appeal.

Appendix

We want to show that, even if the agent notes time's passage, Wagner's result still goes through. This requires two adjustments. First, we want to allow for a change in the agent's credences between q and r, and between q' and r', since her opinions may change as she notices time passing. Second, we don't want to assume that the Jeffrey updates are on the same partitions in opposite orders. If an experience happens first, it will bear directly on a proposition about what things are like at t_0, whereas it will bear directly on a proposition about what things are like at t_1 if it happens second. So we need to consider an update schema of the sort

$$\begin{array}{ccccccc} p & \xrightarrow{\mathbf{E}} & q & \longrightarrow & t & \xrightarrow{\mathbf{F}} & r \\ p & \xrightarrow{\mathbf{F}'} & q' & \longrightarrow & t' & \xrightarrow{\mathbf{E}'} & r' \end{array}$$

where $\mathbf{E}$ is just like $\mathbf{E}'$, except that the propositions in $\mathbf{E}$ are about t_0 and the ones in $\mathbf{E}'$ are about t_1, and analogously for $\mathbf{F}$ and $\mathbf{F}'$. The first and last updates in each case are JC updates on the partitions indicated but, since it's an open question how Bayesians ought to handle awareness of time's passage, I don't want to assume that the t and t' updates happen by JC. But we can assume that t preserves certain probabilistic relationships in q. We are only interested in commutativity for cases where, intuitively, temporal factors aren't relevant to the bearing of the incoming evidence. So let's assume that t doesn't change the conditional probabilities on $\mathbf{F}$ and t' doesn't change the conditional probabilities on $\mathbf{E}'$, so that we have:

$$t(E_i|F_j) = q(E_i|F_j) \text{ for all } i, j, \tag{5}$$

$$t'(F'_j|E'_i) = q'(F'_j|E'_i) \text{ for all } i, j, \tag{6}$$

Thus, while the t and t' updates may not be JC updates, they do not interfere with JC's rigidity.[6] Wagner's proof trades on the rigidity of the updates in his schema, and (5) and (6) will insure that the $q \to r$ and $q' \to r'$ transitions stay rigid with respect to $\mathbf{F}$ and $\mathbf{E}'$ respectively.

Now, we also don't want to assume that $r = r'$ as Wagner does, since we aren't assuming that the order of experience doesn't make any difference. We're just looking at cases where it doesn't make any difference to some hypotheses. In fact, all we have to assume is that the order doesn't make a difference in the agent's credences over the input partitions. More precisely, let's assume that

$$\forall i \forall j : r(E_i F_j) = r'(E'_i F'_j). \tag{7}$$

We also need to assume

$$\forall i_1 \forall i_2 \forall_j : \frac{p(E_{i_2}|F_j)}{p(E_{i_1}|F_j)} = \frac{p(E'_{i_2}|F'_j)}{p(E'_{i_2}|F'_j)} \tag{8}$$

[6] An update from p to q is *rigid* on $\mathbf{E}$ when $q(A|E_i) = p(A|E_i)$ for all i.

$$\forall j_1 \forall j_2 \forall_i : \frac{p(F_{j_2}|E_i)}{p(F_{j_1}|E_i)} = \frac{p(F'_{j_2}|E'_i)}{p(F'_{j_2}|E'_i)} \tag{9}$$

This is another "order doesn't matter" assumption. It says that the initial comparative probabilities of the E-related propositions are the same whether they are about t_0 or t_1, given a corresponding F-related proposition. To get a more definite handle on this, suppose that the E partitions are about a table's color and the F partitions are about its shape. (8) says that the probability that the table is red at t_0 given that it's rectangular at t_1, relative to the probability that it's green at t_0 given that it's rectangular at t_1, is the same as the probability that it's red at t_1 given that it's rectangular at t_0 relative to the probability that it's green at t_1 given that it's rectangular at t_0. In short, the bearing of F-facts on E-facts, even just relatively speaking, is the same whether the E-facts are about t_0 and the F-facts about t_1 or vice versa. And (9) says that the same goes for F-facts given E-facts. These assumption, together with (5) and (6), are enough to run Wagner's proof, as follows.

From (5) and (6), and from the rigidity of JC, we have the identities

$$q(E_i F_j) = q(E_i) p(F_j|E_i) \tag{10}$$

$$q(E_i F_j) = q(F_j) r(E_i|F_j) \tag{11}$$

$$q'(E'_i F'_j) = q'(F'_j) p(E'_i|F'_j) \tag{12}$$

$$q'(E'_i F'_j) = q'(E'_i) r'(F'_j|E'_i) \tag{13}$$

for all i, j. Conjoining (10) with (11) and (12) with (13) yields

$$\forall i \forall j : q(E_i) p(F_j|E_i) = q(F_j) r(E_i|F_j) \tag{14}$$

$$\forall i \forall j : q'(E'_i) r'(F'_j|E'_i) = q'(F'_j) p(E'_i|F'_j) \tag{15}$$

Now if we let $i = i_1$ and then i_2 in (14), and solve for $q(E_{i_1})$ and $q(E_{i_2})$ in each case, we get

$$\beta_{q,p}(E_{i_1} : E_{i_2}) = \frac{p(E_{i_2} F_j) r(E_{i_1} F_j)}{p(E_{i_1} F_j) r(E_{i_2} F_j)} \tag{16}$$

Similar substitutions in (13) yield

$$\beta_{r',q'}(E'_{i_1} : E'_{i_2}) = \frac{p(E'_{i_2} F'_j) r'(E'_{i_1} F'_j)}{p(E'_{i_1} F'_j) r'(E'_{i_2} F'_j)} \tag{17}$$

From (7) and (8) we know that (16) and (17) must be identical, which gives us a bayes factor identity like the one in (3):

$$\forall i \forall j : \beta_{q,p}(E_i : E_j) = \beta_{r',q'}(E'_i : E'_j). \tag{18}$$

Similarly, substituting $j = j_1, j_2$ in (12) and (13) and following analogous reasoning gives us a bayes factor identity like in (4):

$$\forall i \forall j : \beta_{r,q}(F_i : F_j) = \beta_{q',p}(F'_i : F'_j). \tag{19}$$

So the worries raised so far don't eliminate the problem. The order of experience can make a difference but, when it doesn't matter to the input partitions, the same experience has to yield the same bayes factor. Taking the passage of time into account doesn't block this conclusion, so holism is threatened.

We could push harder still, and object that even the modified proof doesn't adequately represent the agent's awareness of temporal considerations. It's not plausible that the input partitions should carry information about what time it is, since the agent may be unsure about the time. Really, experience should bear directly on temporal-indexical propositions like E_i *now*. The agent's beliefs about things like E_i *at* t_0 should then depend on conjoining temporally indexed propositions with her beliefs about what time it is. But, in normal situations, where the agent's beliefs about time are fairly definite and accurate, her beliefs about things like E *at* t_0 will behave as assumed in the proof and the same result will follow.

References

Christensen, David. 1992. "Confirmational Holism and Bayesian Epistemology." *Philosophy of Science* 59.

Field, Hartry. 1978. "A Note on Jeffrey Conditionalization." *Philosophy of Science* 45.

Garber, Daniel. 1980. "Field and Jeffrey Conditionalization." *Philosophy of Science* 47.

Jeffrey, Richard C. 1965. *The Logic of Decision*. University of Chicago Press.

Jeffrey, Richard C. 1975. "Carnap's Empiricism." *Minnesota Studies in the Philosophy of Science* 6.

Lange, Marc. 2000. "Is Jeffrey Conditionalization Defective By Virtue of Being Non-Commutative? Remarks on the Sameness of Sensory Experience." *Synthese* 123.

Oved, Iris and John Pollock. 2005. "Vision, Knowledge, and the Mystery Link." *Philosophical Perspectives* 19.

Skyrms, Brian. 1975. *Choice and Chance*. 2 ed. Dickenson.

Wagner, Carl. 2002. “Probability Kinematics and Commutativity.” *Philosophy of Science* 69.