Conditionalization*

For *The Blackwell Companion to Epistemology, 3rd edition*

eds. Jonathan Dancy, Ernest Sosa, Matthias Steup, and Kurt Sylvan

(*penultimate draft—please cite final version*)

1 The Structure of Conditionalization

Bayesian epistemology's most fundamental diachronic constraint is the norm of Conditionalization. Conditionalization is motivated by the thought that our learning should reflect our past suppositions. If I assign a particular probability to some proposition, supposing that some event occurs, then that probability should be reflected by my doxastic state when I get evidence that this event *has* occurred.

To see how Conditionalization imposes this requirement, suppose that an agent's credence function—the function that represents her degrees of belief—is the probability function, p. We can define p as a function mapping a set of atomic propositions closed under the standard truth-functional connectives to real numbers in the interval [0,1] that satisfy Kolmogorov [1933]'s axioms. The end result is a probability distribution that assigns values to a set of maximally specific, mutually exclusive and jointly exhaustive possibilities, and the propositions associated with the set of possibilities in which these propositions are true. Now suppose that an agent comes to learn E with certainty, and that this represents the strongest proposition that she comes to learn. Then the agent's new credence in any proposition, A, should equal her old credence in A conditional on E, which we represent as the *conditional probability*, $p(A|E)$. Her new probability function, q, should be related to her "prior" probability function, p, according to the following rule:

> **Conditionalization:** If E represents everything that you learn, then for any A, $q(A)=p(A|E)$, if defined.

There are two processes that our credences undergo when we update by Conditionalization. First, all of those possibilities that are incompatible with our evidence receive a credence of 0. Second, all of those possibilities that aren't incompatible with our evidence have their credences changed in a way that preserves the ratios between them, in accordance with Probabilism. These seem like reasonable constraints. If I become certain that it will rain this afternoon, then I shouldn't maintain any credence in possibilities that are incompatible with this information. And those possibilities that are compatible with learning of a rainy afternoon should stand in the same relation to one another as they did before.

Crucial to Conditionalization are the conditional probabilities that encode our past suppositions. It's standard, though not entirely uncontroversial, to define conditional probabilities in terms of unconditional probabilities by means of the *Ratio Formula*:

*Word count: 3299—Including references.

$$p(A|E) = \frac{p(A\&E)}{p(E)}, \text{ if defined.}$$

Drawing on the intuitive connection that holds between updating and confirmation, we can reformulate the Ratio Formula in a way that, as we will see in just a moment, will prove particularly useful. *Bayes' Theorem* asserts the following, for any hypothesis, H, and evidence, E:

$$p(H|E) = \frac{p(E|H)p(H)}{p(E)}, \text{ if defined.}$$

Finally, expanding out the denominator using *the law of total probability* yields the following alternative formulation of Bayes' Theorem:

$$p(H|E) = \frac{p(E|H)p(H)}{\sum_i p(E|H_i)p(H_i)}, \text{ if defined.}$$

It's important to emphasize that unlike Conditionalization, which provides a norm governing the agent's credences at different times, Bayes' Theorem is a mere consequence of the Ratio Formula. What makes it interesting is its ability to express a conditional probability whose value is difficult to discern in terms of quantities whose values are usually more accessible. Consider cases of diagnostic testing, where one knows the likelihoods of testing positive for some disease conditional on the test being accurate ($p(E|H)$) and inaccurate ($p(E|\neg H)$) and, also, the incidence of this disease in the population $p(H)$. In such cases, it's easy to determine the probability of being sick conditional on a positive test result ($p(H|E)$).

Since experiments often begin with well-defined priors in competing hypotheses and likelihoods, Bayes' Theorem provides us with the values that we need to update by Conditionalization. It's in virtue of this that those who revise their beliefs according to this rule have come to be called "Bayesians".

2 Jeffrey Conditionalization

Since Conditionalization is triggered by getting evidence, it implies the existence of an experience that delivers this evidence by changing the agent's credence in some proposition to one. However, an account that acknowledges that our updates begin in experience suggests that our evidence will sometimes be uncertain. For we often experience the world in less than favorable conditions, as when we observe a cloth in dim lighting that may be green or blue or possibly even purple.

These types of examples led Jeffrey [1965] to propose a generalization of Conditionalization that takes our evidence to be a *partition* of propositions (a set of mutually exclusive and jointly exhaustive propositions) with a set of values that sum to one:

> **Jeffrey Conditionalization:** When an experience directly changes the probabilities over a partition, $\{E_i\}$, from $p(E_i)$ to $q(E_i)$, the new probability for

any proposition A, should be

$$q(A) = \sum_i p\left(A \mid E_i\right) q\left(E_i\right), \text{if defined.}$$

While Jeffrey Conditionalization is a formal generalization of "Strict" Conditionalization, it raises a number of epistemological concerns. Once we abandon the constraint that our evidence be a proposition that we learn with certainty, we seem to be left with something less than a diachronic constraint. Unless we adopt some account of evidence, there is nothing other than the probability axioms to constrain the evidence an agent updates on, which makes Jeffrey Conditionalization look less like a diachronic norm, and more like a description of the way such an agent satisfies Probabilism over time. Perhaps the easiest way of appreciating this concern about the lack of an evidential constraint is to notice that, unlike Strict Conditionalization, Jeffrey Conditionalization fails to satisfy the very minimal condition that our evidence be order-invariant—that it be *commutative*.

Some have claimed that these problems, especially the non-commutativity of Jeffrey Conditionalization, suggest that the Jeffrey framework assumes a different picture of the formal structure of our evidence. Field [1978] proposes a reparametrized version of Jeffrey Conditionalization that takes the inputs to the framework to be phenomenal experiences, which are formally represented as the *Bayes factors* that encode the magnitudes of the changes over the evidence partitions of the updates in question. As Lange [2000] points out, the fact that Jeffrey Conditionalization isn't commutative over weighted evidence partitions doesn't entail that it isn't commutative over these experiences. However, both Christensen [1992] and Weisberg [2009] note that taking experiences to be the inputs to the Bayesian framework preserves the commutativity of Jeffrey Conditionalization at the cost of making it anti-holistic. It does this by preventing our background beliefs from mediating the impact that our experiences have on our updates.

We've noted already that fundamental to Conditionalization is the idea that our updates should be guided by our conditional probabilities. By focusing on the proportional shifts among the propositions in our evidence partitions, Jeffrey Conditionalization points us towards an equivalent principle that makes this fundamental idea even more transparent. Where p and q, again, represent the agent's old and new credence functions, respectively, the *Rigidity Principle* imposes the following constraint on any arbitrary proposition, A, for each proposition in the evidence partition, $\{E_i\}$:

$$p(A|E_i) = q(A|E_i), \text{ if defined.}$$

The *Rigidity Principle* highlights the guiding feature of our conditional probabilities by requiring that those conditional probabilities that are relevant to an update remain "rigid" or unchanged over this update.

3 Rational Constraints on Priors

Many Bayesians endorse the more permissive account of evidence that Jeffrey Conditionalization assumes. This section considers a different way that Conditionalization can be made more or less permissive.

While all Bayesians maintain that an agent's prior credence function ought to satisfy Probabilism, Subjective Bayesians assume that this function is at least partially unconstrained. They hold that there are a range of permissible prior probability distributions that encode the relations of evidential support we deploy when we update by Conditionalization. By contrast, Objective Bayesians maintain that there is never more than one attitude it is rational to have in response to one's total evidence. They hold that there is a uniquely rational prior probability distribution.

The constraint that is most closely associated with Objective Bayesianism is the *Principle of Indifference*:

> Given a set of possibilities that form a partition, you should assign the same probability to each possibility in the absence of any evidence favoring one over any other.

While the Principle of Indifference aims to provide a single constraint that determines a unique prior probability distribution, it famously yields inconsistent prescriptions since its values depend upon how we partition the space of possibilities (*see* van Fraassen [1989] for the classic illustration of this). A different principle some have thought ought to constrain an agent's entire credence distribution is the *Regularity Principle*:

> No logically contingent proposition should receive credence 0.

The Regularity Principle encodes the plausible thought that we should never entirely rule out any logical possibility. Of course, where we take this principle to hold at every stage of inquiry, this plausible thought conflicts with Strict Conditionalization.

Other constraints govern local parts of an agent's prior probability distribution. Many of these constraints can be broadly classified as "deference principles" since they tell us that we should defer to the opinions of experts. Two of the most well-known deference principles are the *Reflection Principle* and the *Principal Principle*.

Van Fraassen [1984]'s *Reflection Principle* is motivated by the idea that one should defer to one's future opinion. If I know that tomorrow I'll have a credence of .9 in the proposition that a Democrat will win the election, then I should maintain that credence in that proposition today. For any arbitrary proposition, A, our current credences, cr_t ought to align with what we take our future credences, $cr_{t'}$, to recommend in roughly the following way:

$$cr_t(A \mid cr_{t'}(A)=\mathrm{x})=\mathrm{x}$$

The Reflection Principle runs into problems in cases where one expects one's future judgment to be deficient, due to memory loss, impairment or any number of other cognitive shortcomings. Amendments to this principle are needed to circumvent these problems, some of which have been proposed.

A different type of expert comes in the form of objective probabilities. Lewis [1980]'s *Principal Principle* takes our knowledge of the objective chances to impose a constraint on our credences. Where A is some arbitrary proposition, cr is some reasonable initial credence function, and ch_t is the chance function at time, t, the *Principal Principle* recommends something roughly along the following lines:

$$cr(A \mid ch_t(A) = x) = x$$

This simplified version of the Principal Principle does not hold unrestrictedly. It fails to apply in cases where we have so-called inadmissible information: information that does not affect what probability we think we ought to assign to some proposition by way of affecting the objective chance of this proposition. While Lewis does not provide a precise definition of admissible information, he took it to include information about the history of the world before t, as well as information about how chances depend upon that history. As with the Reflection Principle, there have been numerous proposals for formulating a version of the Principal Principle that best captures these, and other, restrictions (*see* Meacham [2010]).

4 Justifying Conditionalization

While Conditionalization is an intuitively attractive updating rule, one might ask whether more can be said in its defense.

Historically, the justification for Bayesianism has its source in the observation that degrees of belief that fail to satisfy the probability axioms are associated with betting quotients that can be exploited by a clever bookie to produce a sure loss (*see* Ramsey [1931], de Finetti [1937]). Such a collection of bets is called a *Dutch Book*. The Dutch Book argument for Conditionalization draws upon the *Dutch Strategy* developed by David Lewis and reported in Teller [1973]. What differentiates a Dutch Strategy from a Dutch Book is that, rather than involving a fixed set of bets, a Dutch Strategy is a strategy for placing certain bets at different times that the agent would regard as fair, where some of these bets will depend upon what the agent learns. The Dutch Book argument for Conditionalization begins with the idea that failing to update by Conditionalization leaves the agent open to a Dutch Strategy—and, thus, to a sure loss of money—and concludes that we ought to update by Conditionalization.

The idea that we ought to avoid Dutch Strategies has been criticized on a number of grounds. One problem is that this idea assumes that a series of bets that are regarded as fair when considered individually ought to also be regarded as fair when considered as a package. This seems especially dubious given that, in the case of Dutch Strategies, these bets are being evaluated at two different times, relative to two different sets of probabilities. Another widely recognized problem with the traditional Dutch Book argument for Conditionalization is that the strongest norm it establishes isn't the diachronic norm of Conditionalization, but the synchronic norm to plan to conditionalize on a particular partition.

A different criticism that has been raised against all Dutch Book arguments is that they provide practical rather than epistemic reason to conform to epistemic norms. Accuracy arguments overcome this challenge by grounding their justifications in the epistemic value of gradational accuracy. Greaves and Wallace [2006], and others, offer accuracy arguments that justify the norm to plan to conditionalize. Others offer accuracy arguments that claim to justify a genuinely diachronic norm of Conditionalization. While a comprehensive examination of these arguments cannot be undertaken in this discussion, it's worth noting that accuracy arguments rely upon much of the same mathematical framework that Dutch Book arguments assume.

5 Memory Loss, Old Evidence, and Context-Sensitivity

Finally, it's worth considering three problems for Conditionalization. The first is that Conditionalization does not allow the agent to lose evidence that she has gained with certainty. If a proposition has been assigned a credence of one, then the agent's credence in that proposition, conditional on any other proposition, must be one as well. Therefore, once one maintains a credence of one in some proposition, one's credence in that proposition cannot drop as the result of any future learning episode. This is problematic since it makes memory loss impossible if we assume Strict Conditionalization.

A second problem that many take to follow from the certainty of evidence is *the problem of old evidence*, described by Glymour [1980]. Consider the case where we learn that some old evidence, E, is entailed by some hypothesis, H. It's natural to think, in this case, that this evidence confirms this hypothesis, so that $p(H|E)>p(H)$. But where E is evidence we have already learned, then $p(E)=1$, and, so, $p(H|E)=p(H)$. Thus, given a natural assumption about the relation between updating and confirmation—that one's hypothesis H is confirmed by E exactly when one's credence in H ought to increase in response to learning that E—we are unable to discover that old evidence confirms some theory that we hold.

One way of avoiding these difficulties is to adopt Regularity (*see* §3) and Jeffrey Conditionalization (*see* §2). A different approach to the problem of losing certainties is to appeal to a prior probability distribution that does not encode a memory. Rather than maintain that an agent's prior function is the probability function that she had before the last time that she updated, we might appeal to an initial or "*ur*-prior" distribution, which is the function that we assume an agent to have had before she learned anything whatsoever. Since a *ur*-prior function takes as input the total evidence an agent has at the time that she updates, it is able to deliver the right result in cases where the agent has lost evidence that she previously held with certainty.

While many solutions have been proposed to the problems of memory loss and old evidence, it's worth noting that not everyone takes these problems to be fundamentally about the certainty of evidence. Christensen [1999] argues that a more general version of the problem of old evidence persists even if we assume that our evidence is less than completely certain. He ends up concluding that no Bayesian measure is able to capture all of our intuitive judgments about confirmation (though *see* Eells and Fitelson [2000] for a rejoinder). In the same spirit, Hedden [2015, p.38-39] suggests that the more general problem raised in certain cases of potential memory loss persists even for updates on uncertain evidence.

A final challenge for Conditionalization are situations that rationally require us to update self-locating beliefs. Among the beliefs I have are beliefs, not just about what the world is like, but beliefs about my place in it—for instance, my belief that "Today is Thursday". I might have some evidence that leaves me certain of that claim today. But I clearly shouldn't be certain of it tomorrow. Conditionalization entails that I must be.

It's tempting to think that the previous problem is merely an instance of the problem of being unable to lose certain evidence. Part of the worry in the case just described is that I'm committed to my belief that "Today is Thursday" on Friday. But even if I weren't committed on Friday to the belief that "Today is Thursday", we would still need more fine-grained contents—"centered propositions"—to model such beliefs in the first place, and to revise them so as to take into account the systematic way that context-sensitive information shifts over time. A number of alternatives to Conditionalization, defined over these more fine-grained contents, that attempt to get us the right dynamics for self-locating beliefs, have been proposed (*see* Titelbaum [2016] for a taxonomy of the different strategies).

References

David Christensen. Confirmational holism and bayesian epistemology. *Philosophy of Science*, 59(4):540–557, 1992.

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

Bruno de Finetti. La prévision: ses lois logiques, ses sources subjectives. In *Annales de l'institut Henri Poincaré*, volume 7, pages 1–68, 1937.

Ellery Eells and Branden Fitelson. Measuring confirmation and evidence. *The Journal of philosophy*, 97(12):663–672, 2000.

Hartry Field. A note on jeffrey conditionalization. *Philosophy of Science*, 45(3): 361–367, 1978.

Clark Glymour. *Theory and Evidence*. Princeton University Press, 1980.

Hilary Greaves and David Wallace. Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind*, 115(459):607–632, 2006.

Brian Hedden. *Reasons without persons: rationality, identity, and time*. Oxford University Press, 2015.

Richard C Jeffrey. *The Logic of Decision*. McGraw-Hill Book Co Inc. New York, 1965.

Andrei N Kolmogorov. *Foundations of the theory of probability*. Chelsea Publishing Company, 1933.

Marc Lange. Is jeffrey conditionalization defective by virtue of being non-commutative? remarks on the sameness of sensory experiences. *Synthese*, 123 (3):393–403, 2000.

David Lewis. A subjectivist's guide to objective chance. In *Ifs*, pages 267–297. Springer, 1980.

Christopher JG Meacham. Two mistakes regarding the principal principle. *British Journal for the Philosophy of Science*, 61(2):407–431, 2010.

Frank P Ramsey. Truth and probability. In R. B. Braithwaite, editor, *The foundations of mathematics and other logic essays*, pages 156–198. New York: Harcourt, Brace and Company, 1931.

Paul Teller. Conditionalization and observation. *Synthese*, 26(2):218–258, 1973.

Michael G. Titelbaum. Self-locating credences. In Alan Hajek and Christopher Hitchcock, editors, *The Oxford Handbook of Probability and Philosophy*. Oxford University Press, 2016.

Bas C Van Fraassen. Belief and the will. *The Journal of Philosophy*, 81(5):235–256, 1984.

Bas C van Fraassen. *Laws and symmetry*. Clarendon Press, 1989.

Jonathan Weisberg. Commutativity or holism? a dilemma for conditionalizers. *British Journal for the Philosophy of Science*, 60(4):793–812, 2009.